

Ansgar Franz Breitung

MANAGEMENT-INFORMATIONSSYSTEME

**Ein methodologischer Beitrag zur Analyse
von Datenverarbeitungsprozessen
in Marketing-Informationssystemen**

1973

1974

Management-Informationssysteme:
Ein methodologischer Beitrag zur Ana-
lyse von Datenverarbeitungsprozessen in
Marketing-Informationssystemen

Inaugural-Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
durch den Fachbereich
Wirtschafts- und Sozialwissenschaften
der Westfälischen Wilhelms-Universität
zu Münster

vorgelegt von
Ansgar Franz Breitung
aus Quakenbrück
1974



Erster Berichterstatter: Prof. Dr. H. Meffert

Zweiter Berichterstatter: Prof. Dr. H. Wagner

Dekan: Prof. Dr. D. Adam

Tag der mündlichen Prüfung: 20. November 1973

Inhaltsverzeichnis

	Seite
<u>I. Einleitung und Problemstellung</u>	1
1. Begriff und Elemente eines Marketing-Informationssystem	1
2. Forschungsansätze bei Marketing-Informationssystemen	12
3. Problemstellung und Gang der Untersuchung	18
 <u>II. Die Datenbank im Marketing-Informationssystem</u>	 23
1. Datenbankbegriff und Datenbanktypen	23
2. Organisation der Datenbank	25
2.1 Formale Datenorganisation	25
2.1.1 Sprachanalytische Segmentierung	25
2.1.2 Logische Segmentierung	42
2.2 Physische Datenorganisation	51
2.2.1 Indizierung von Datenelementen	51
2.2.2 Speicherung von Datenelementen	64
2.3 Technisch-physische Datenorganisation	67
2.3.1 Zugriffszeiten	67
2.3.2 Zugriffswege	71
3. Verwaltung der Datenbank	74
3.1 Kommunikation mit der Datenbank	74
3.1.1 Grundprobleme der Datenbank-Kommunikation	74
3.1.2 Direktanfragen an die Datenbank	79
3.2 Anpassung der Datenbank	83
3.3 Schutz der Datenbank	89
4. Softwareentwicklungen für Datenbanken	95

	Seite
4.1 Grundproblem der Datenbank-Software	95
4.2 Integrated Data Store	97
 III. <u>Die Methodenbank im Marketing-Informationssystem</u>	 106
1. Methoden als Komponenten der Methodenbank	106
1.1 Zum Begriff der Methode	106
1.2 Inhalt und Aufgaben der Methodenbank	108
2. Einordnung multivariater Daten-Analysemethoden	113
3. Verfahren zur Analyse von Abhängigkeiten	118
3.1 Regressionsanalyse	118
3.2 Varianzanalyse	125
3.3 Kovarianzanalyse	138
4. Klassifikations- und Typisierungsverfahren	146
4.1 Diskriminanzanalyse	146
4.2 Clusteranalyse	152
5. Verfahren zur Analyse von Zusammenhängen	159
5.1 Korrelationsanalyse	159
5.2 Faktorenanalyse	163
 IV. <u>Die Modellbank im Marketing-Informationssystem</u>	 179
1. Modelle als Komponenten der Modellbank	179
1.1 Modellarten	179
1.2 Funktionen und Inhalt der Modellbank	186
2. Ansätze zur Modellselektion und-entwicklung	193
2.1 Der theoretisch-intersubjektive Ansatz	193
2.2 Der benutzerorientierte-subjektive Ansatz	210

	Seite
3. Analyse von Kriterien zur Modellselektion und -entwicklung	229
3.1 Einfachheit und Prüfbarkeit	229
3.2 Prognostische Relevanz	236
3.3 Adaptivität und Kommunikation	251
3.4 Berechenbarkeit und Vollständigkeit	255
3.5 Integrationsfähigkeit	269
4. Schlußfolgerungen für die Modellselektion und -entwicklung	283
 V. <u>Ausblick auf künftig mögliche Forschungsstrategien bei Marketing-Informationssystemen</u>	 292
 Literaturverzeichnis	 308

Abbildungsverzeichnis

	Seite
Abb. 1: Funktional abgegrenztes Informationssystem	2
Abb. 2: Struktur und Elemente eines Marketing-Informationssystems	10
Abb. 3: Sprachanalytische Segmentierung	40
Abb. 4: Logische Ordnungskriterien	45
Abb. 5: Logischer Baum	47
Abb. 6: Verkettete Organisationsform	58
Abb. 7: Zugriffsstufenproblem	62
Abb. 8: Kette im Integrated Data Store	98
Abb. 9: Kettung dreier hierarchischer Strukturen im Integrated Data Store	100
Abb. 10: Zugriffsweg beim Integrated Data Store	104
Abb. 11: Systematik multivariater Daten-Analysemethoden	116
Abb. 12: Berechnung der Entfernung zwischen Kombinationen von zwei Merkmalsausprägungen	154
Abb. 13: Clusterbildung bei zwei Merkmalsausprägungen	157
Abb. 14: Orthogonale Faktorrotation	175
Abb. 15: Zusammenhang zwischen abzubildendem Original, Benutzer und Modell	197
Abb. 16: Modell- und Datenfehler	205
Abb. 17: Bestimmung des Minimums des Gesamtfehlers	209
Abb. 18: Problemlösungsprozeß	221
Abb. 19: Integrationsbedingungen von Prognosemodellen	275
Abb. 20: Begreifen von realen Objekten	294

Tabellenverzeichnis

	Seite
Tab. 1: Ausgangsdaten der einfachen Varianzanalyse	126
Tab. 2: Ausgangsdaten der zweifachen Varianzanalyse	132
Tab. 3: Ausgangsdaten der einfachen Kovarianzanalyse	140
Tab. 4: Ausgangsdaten zur Bestimmung der Diskriminanzfunktion	147
Tab. 5: Metrisch skalierte Merkmalsausprägungen der Käufer bei zwei Variablen	153
Tab. 6: Absolute Altersunterschiede zwischen den Käufern	153
Tab. 7: Absolute Einkommensunterschiede zwischen den Käufern	154
Tab. 8: Euklidische Entfernungen bei zwei Merkmalsausprägungen	155
Tab. 9: Berechnung der Entfernungen zwischen dem Schwerpunkt und Merkmalsausprägungen	156
Tab. 10: Charakteristik von analytischen, quasi-analytischen und heuristischen Modellen	184
Tab. 11: Einstufung von Problemlösungsmethoden durch Relevanzziffern von 1 bis 10	190

Abkürzungsverzeichnis

a.a.O.	am angegebenen Orte
Abb.	Abbildung
Aufl.	Auflage
bzw.	beziehungsweise
d.h.	das heißt
Diss.	Dissertation
et al.	et alteri
etc.	et cetera
f	folgende Seite
ff	folgende Seiten
ggf.	gegebenenfalls
H.	Heft
Hrsg.	Herausgeber
hrsg. v.	herausgegeben von
i.d.R.	in der Regel
i.e.S.	im engeren Sinn
i.w.S.	im weiteren Sinn
Jg.	Jahrgang
No.	Number
Nr.	Nummer
o.a.	oben angeführt
o.J.	ohne Jahr
S.	Seite
u.a.	unter anderem
usw.	und so weiter
vgl.	vergleiche
Vol.	Volume
z.B.	zum Beispiel

I. Einleitung und Problemstellung

1. Begriff und Elemente eines Marketing-Informationssystems

Das übergeordnete Bezugssystem für Forschungen auf dem Gebiet von Informationssystemen ist die Unternehmung in ihrer Gesamtheit. Unternehmungen sind in diesem Zusammenhang offene, zielgerichtete sozio-technische Entscheidungssysteme, die aus einer Menge von Elementen - Menschen und Maschinen - bestehen, welche in stofflich-energetischem und informationellem Kontext zur Umwelt stehen.¹ Ein sozio-technisches System ist definiert als "geordnetes Zusammenwirken von Menschen und Technik zur Erreichung eines Zielles."² Das sozio-technische System "Unternehmung" kann funktional strukturiert werden, wobei diese Strukturierung in gewissen Grenzen immer eine Frage der Zweckmäßigkeit und infolgedessen nicht zwingend ist.

Bei der Definition des Begriffs "Informationssystem" lassen sich aufgrund des übergeordneten Bezugssystems zwei Auffassungsrichtungen unterscheiden. Nach der ersten Auffassungsrichtung i s t die Unternehmung ein Informationssystem. Sämtliche realen Prozesse finden ihre Entsprechung in Informationen, die es ermöglichen, diese Prozesse zweckorientiert zu regeln und zu steuern. Die zweite Auffassungsrichtung begreift das Informationssystem als funktionales Subsystem der Unternehmung, das durch spezifische Verhaltensweisen gekennzeichnet ist.

Die folgende Untersuchung geht von dieser zweiten Auffassungsrichtung aus. Danach h a t die Unternehmung ein Informationssystem. Es wird durch die Identifizierung seiner Elemente und Relationen zu anderen Subsystemen der Unternehmung abgegrenzt.

1 Vgl. Meffert, H., Zum Problem der betriebswirtschaftlichen Flexibilität, in: Zeitschrift für Betriebswirtschaft, 39. Jg., 1969, S. 779ff, hier S. 781.

2 Wedekind, H., Datenorganisation, Berlin 1970, S. 232.

Elemente dieses Verhaltenssystems sind Organisationsteilnehmer, die entsprechend dem Prinzip der Partialinklusion in mehreren funktionalen Subsystemen spezifische Rollen zu erfüllen haben.¹ Dieses funktional abgegrenzte Informationssystem ist in seinem unternehmensbezogenen Kontext in der folgenden Abbildung dargestellt.

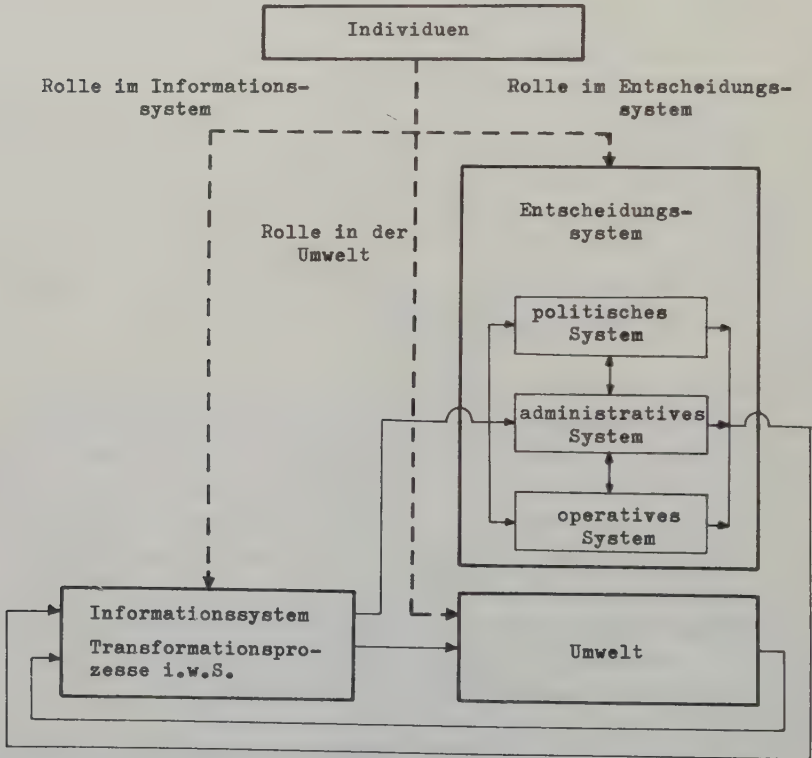


Abb. 1: Funktional abgegrenztes Informationssystem²

¹ Vgl. Katz, D. / Kahn, R.L., The Social Psychology of Organizations, New York usw. 1967, S. 50ff und Meffert, H., Systemtheorie aus betriebswirtschaftlicher Sicht, in: Wirtschaftskybernetik und Systemanalyse, hrsg. v. K.-E. Schenk, Berlin 1971, S. 174ff, hier S. 184.

² In Anlehnung an Meffert, H., Informationssysteme - EDV und Systemanalyse, Teil I: Grundlagen, Münster 1973, S. 52.

Das Informationssystem gewinnt aus der Umwelt und dem Entscheidungssystem, das als Zusammenfassung des politischen, administrativen und operativen Systems verstanden wird, informationistischen Input. Dieser wird in Output transformiert, der an der Umwelt und an betreffende Subsysteme des Entscheidungssystems abgegeben wird. Darüber hinaus sind aus der Abbildung die Rollen ersichtlich, die die Individuen im Informationssystem, in der Umwelt und im Entscheidungssystem zu erfüllen haben. Die Aufgaben der funktionalen Subsysteme des Entscheidungssystems gilt es im folgenden unter Beachtung der informationsgewinnenden und -verarbeitenden Funktionen des Informationssystems zu präzisieren.

Das politische System¹ hat die Aufgabe, konfliktäre Zielvorstellungen seiner Elemente zum Ausgleich zu bringen. Es verarbeitet Informationen, die vom Informationssystem aus der Umwelt oder aus nachgelagerten funktionalen organisatorischen Subsystemen gewonnen werden und leitet diese an das administrative oder operative Subsystem weiter. Die in das politische System eingehenden Inputs werden in Führungsentscheidungen transformiert und an nachgelagerte Subsysteme weitergeleitet. Die Zielformulierung als originärer Output des politischen Systems ist dabei von den informationellen Inputs abhängig, die es vom Informationssystem erhält.

¹ Zur Aufgabenstellung des politischen Systems vgl. Kutschker, M., Verhandlungen als Elemente eines verhaltenswissenschaftlichen Bezugsrahmens des Investitionsgütermarketing, Diss., Mannheim 1972, S. 9f und Fäßler, K., Betriebliche Mitbestimmung - Verhaltenswissenschaftliche Projektionsmodelle, Bd. 6 der Schriftenreihe "Die Betriebswirtschaft in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1970, S. 87ff.

Während dem politischen System unter Mitwirkung des Informationssystems primär die Aufgabe zukommt, die Richtlinien der langfristigen Unternehmenspolitik festzulegen, hat das administrative System darüber zu entscheiden, wie die vom politischen System formulierten Ziele und Strategien verwirklicht werden können.¹

Schließlich hat das operative System die originäre Aufgabe, Vollzugsentscheidungen zu fällen und zu realisieren. Diese Entscheidungen basieren auf Programmen, die vom administrativen System erarbeitet werden.

Im Rahmen einer derart skizzierten Aufgabenstellung des Informationssystems ist ersichtlich, daß dadurch sämtliche Funktionsbereiche einer Unternehmung betroffen werden. Aus diesem Grund wird oft der Schluß gezogen, daß ein Informationssystem sämtliche Funktionsbereiche einer Unternehmung umfassen soll (integriertes Informationssystem)², die es mit Informationen zu versorgen hat. Ihrem Wesen nach sind Führungsentscheidungen stets integrativ, so daß es gerechtfertigt ist, den Begriff des Management-Informationssystems mit dem des integrierten Informationssystems gleichzusetzen.

Problem ist, inwieweit sich eine konsistente und sinnvolle Abgrenzung zwischen einem Management- und einem Marketing-Informationssystem finden läßt. Eine diesbezügliche Abgrenzung ist notwendig, weil in jüngster Zeit Vorschläge über den Aufbau von eigenständigen Marketing-Informationssystemen erbracht worden sind.³

1 Vgl. dazu Kirsch, W., Entscheidungsprozesse, Bd. III: Entscheidungen in Organisationen, Wiesbaden 1971, S. 121ff.

2 Vgl. Grochla, E., Forschung und Entwicklung auf dem Gebiet der Informationssysteme als Aufgabe der Betriebswirtschaftslehre, in: Zeitschrift für Betriebswirtschaft, 41 Jg., 1971, S. 563ff.

3 Vgl. dazu Köhler, R., Informationssysteme für die Unternehmensführung, in: Zeitschrift für Betriebswirtschaft, 41. Jg., 1971, S. 27ff und Kotler, P., Die Zukunft des Computers im Marketing, in: Der Markt, 1970, S. 81ff.

Marketing wird hier als Planung, Koordination und Kontrolle aller auf die aktuellen und potentiellen Märkte ausgerichteten Unternehmensaktivitäten verstanden, die einerseits auf die Erreichung der Unternehmensziele, andererseits auf die Befriedigung der Kundenbedürfnisse ausgerichtet sind.¹ Marketing zeichnet sich besonders dadurch aus, daß eine informationsgesteuerte Transaktion von Real- und Nominalgütern vom Absatzmarkt her vorzunehmen ist. Diese Aufgabe ist um so eher gewährleistet, wenn es gelingt, die dazu notwendigen Informationen jederzeit den betreffenden Entscheidungsträgern zugänglich zu machen. Aufgrund der beschränkten menschlichen Informationsverarbeitungskapazitäten und der relativ langen Zeitdauer, die für die Weiterleitung von Informationen notwendig ist, ergibt sich die Konsequenz, daß ein Marketing-Informationssystem computergestützt sein muß.²

Darüber hinaus ist es nach bestimmten Prinzipien strukturiert und generiert unter Bezugnahme auf bestimmte Methoden und Modelle Informationen für Marketing-Problemstellungen. Die Gesichtspunkte der interaktiven Computerstützung - d.h. das Zusammenwirken von Computer und Benutzer - und der Generierung von Informationen für Marketing-Problemstellungen mit Hilfe von Methoden und Modellen gilt es zu analysieren, um sowohl den Begriff des Marketing-Informationssystems zu präzisieren als auch eine Abgrenzung vom Management-Informationssystem vorzunehmen.

MONTGOMERY³ sieht das Abgrenzungsproblem zwischen Management-Informationssystem und Marketing-Informationssystem

1 Vgl. Meffert, H., Art.: Marketing, in: Management-Enzyklopädie, Bd. IV, München 1971, S. 383ff, hier S. 383.

2 Vgl. u.a. Cox, D.E. / Good, R.E., How to build a Marketing Information System, in: Harvard Business Review, May-June 1967, S. 145ff; Montgomery, D.B. / Urban, G.L., Marketing Decision-Information Systems: An Emerging View, in: Journal of Marketing Research, May 1970, S. 226ff.

3 Vgl. Montgomery, D.B., Marketing Information and Decision Systems: Coming of Age in the '70's, Graduate School of Business - Stanford University, Research Paper No. 150, April 1973, S. 2.

primär aus terminologischer Warte: "Um nicht von der allgemeinen Betriebswirtschaftslehre überholt zu werden, eignete sich der Marketing-Bereich ... diese Begriffsfolge [MIS = Management-Informationssystem] an und benannte sie mit Marketing-Informationssystem." Im folgenden soll eine präzisere Abgrenzung vorgenommen werden, die 1. auf dem Systembenutzer basiert, der sich 2. Methoden und Modelle für Marketing-Problemstellungen bedient, die 3. mit aus der Umwelt und dem Entscheidungssystem gewonnenen Daten arbeiten.

Wird der Systembenutzer zur Abgrenzung zwischen Management- und Marketing-Informationssystem herangezogen, so entsteht das Problem, welche Organisationsteilnehmer als Elemente des Management-Informationssystems zu betrachten sind. Geht man vom funktional abgegrenzten Entscheidungssystem in Abbildung 1 aus¹, so ist bei den Entscheidungen seiner Subsysteme ein abgestufter Zusammenhang zum Gesamtunternehmensgeschehen entsprechend ihrer Rangstufe festzustellen. D.h., Entscheidungen des politischen Systems müssen den Interdependenzzusammenhang stärker beachten als Entscheidungen des operativen Systems. Daraus folgt, daß diejenigen Organisationsteilnehmer als Elemente eines Management-Informationssystems zu betrachten sind, die Führungsentscheidungen zu treffen haben. Marketing-Manager wären somit aufgrund des Interdependenzzusammenhangs und der von Ihnen zu fällenden Entscheidungen Elemente eines Management-Informationssystems.

Die Problemstellungen als zweiter Abgrenzungsfaktor zwischen Management- und Marketing-Informationssystem, die mittels Methoden und Modellen im Marketing gelöst werden sollen, werden von den Marketing-Organisationsteilnehmern spezifiziert. Methoden und Modelle lassen sich aber nur dann anwenden, wenn ihre Inputdaten beschafft worden sind. Inputdaten von Methoden und Modellen ergeben sich aus der Umwelt und

¹ Vgl. S. 2 dieser Arbeit.

dem Entscheidungssystem. Aus Wirtschaftlichkeitsgründen wäre es wenig sinnvoll, auf Methoden und Modelle zu verzichten, die auf Marketingdaten (Umsatz, Marktanteil etc.) und Daten anderer Funktionsbereiche der Unternehmung angewiesen sind und die für diese Funktionsbereiche wesentliche Ergebnisse liefern.

Hieraus ergibt sich für die in ein Management-Informationssystem eingehenden Inputdaten der Schluß, daß diese nur insoweit für eine Abgrenzung zwischen Management-Informationssystem und Marketing-Informationssystem herangezogen werden können als sie marketingspezifisch sind. Da eine diesbezügliche Klassifikation aber oft auf Schwierigkeiten stoßen wird, ist es unter Beachtung des Zusammenhangs im Unternehmen sinnvoll, Management-Informationssysteme aufzubauen.

Es ergibt sich die Konsequenz, daß ein Marketing-Informationssystem als "Menge" von Maschinen, Marketing-Organisationsteilnehmern, marketingrelevanten Methoden und Modellen und marketingspezifischen Daten innerhalb eines Management-Informationssystems interpretiert werden kann. Faßt man die im Rahmen eines Management-Informationssystems verwendeten Methoden in einer Methodenbank, die Modelle in einer Modellbank und die Daten in einer Datenbank zusammen, so läßt sich die Abgrenzungproblematik zwischen Management-Informationssystem und Marketing-Informationssystem wie folgt charakterisieren: Unter dem Gesichtspunkt, daß alle Funktionsbereiche einer Unternehmung mit Informationen versorgt werden sollen, müssen vom integrativen Ansatz her Management-Informationssysteme entwickelt werden. Die Grenze zwischen beiden Systemen ist nur so zu ziehen, daß eine gedankliche Trennung zwischen Methoden, Modellen und Daten vorgenommen wird, die marketingspezifisch sind. Daraus folgt, daß ein Management-Informationssystem durchaus funktional ein Marketing-Informationssystem sein kann, wenn es in erster Linie bei den Methoden, Modellen und Daten auf Marketingzwecke hin orientiert ist.

Der Grund, daß eine diesbezügliche Betrachtungsweise in der Regel nicht erfolgt, ist aufbaustrategisch zu erklären. Er liegt in dem extrem problematischen Unterfangen, sofort mit dem Aufbau von Management-Informationssystemen zu beginnen. Zum gegenwärtigen Zeitpunkt gehören viele Überlegungen, die auf dem Gebiet der Management-Informationssysteme angestellt werden, angesichts mannigfaltiger Schwierigkeiten - wie z.B. der Lösung des Integrationsproblems - in den Bereich der Utopie.

Für diese Schwierigkeiten sind u.a. folgende Gründe ausschlaggebend: Die für einen integrativen Ansatz notwendigen Daten sind fast nie in den Formen vorhanden, wie sie benötigt würden. Darüber hinaus erweist sich der Zugriff zu den Datenbeständen auf Grund der Art und Weise wie diese Daten zu organisieren und zu speichern sind als informationstechnologisch problematisch.¹ Ein weiterer Grund, von der Entwicklung von Management-Informationssystemen Abstand zu nehmen, ist in der Entscheidungsstruktur des Managements begründet. Integrative Entscheidungen, die insbesondere durch Interdependenzen mit mehreren Funktionsbereichen einer Unternehmung gekennzeichnet sind, erfolgen nur selten routinemäßig. Darüber hinaus zeichnen sich bislang Modelle integrativer Unternehmensprozesse durch ihre mangelnde Praktikabilität aus, so daß es nur bedingt möglich ist, integrative Entscheidungen aufgrund von informationellen Outputs eines Management-Informationssystems zu verbessern. Das logisch sinnvolle Ziel - die Entwicklung von Management-Informationssystemen - muß zugunsten der Entwicklung von Informationssystemen auf spezifischen unternehmerischen Funktionsebenen aufgegeben werden. Diese Informationssysteme müssen im Weg der "Politik der kleinen Schritte" aufgebaut werden.

Die hier nur umrißhaft aufgezeigten Probleme sind beim Aufbau von Informationssystemen auf Subsystemebene - so beim Marketing-Informationssystem - weniger komplex. Die-

¹ Vgl. US Department of Housing and Urban Development (HUD), Urban and Regional Information Systems (URIS) - Support for Planning in Metropolitan, Teil 1, Washington D.C. 1968, S.27.

ses ist darin begründet, daß der Interdependenzzusammenhang bei Marketing-Informationssystemen nicht in dem Ausmaß wie bei Management-Informationssystemen beachtet zu werden braucht. Die primäre Zwecksetzung, die mit einem diesbezüglichen Informationssystem erreicht werden soll, bestimmt somit dessen Nomenklatur.

Ausgehend von diesem Sachverhalt, kann ein Marketing-Informationssystem wie folgt definiert werden: Es ist ein funktionales auf Marketingzwecke hin orientiertes, nach bestimmten datenverarbeitungsspezifischen Gesichtspunkten strukturiertes Mensch-Maschine-System, das auf Basis kommunikativer Beziehungen in der Marketing-Organisation beruht, um Problemlösungen mittels Methoden und Modellen zu generieren, die mit Daten aus einer Datenbank arbeiten, wobei Daten, Methoden und Modellergebnisse an seine Benutzer über eine Ein- und Ausgabeeinheit übermittelt werden.

Die Komponenten eines solchen Marketing-Informationssystems sind in Abb. 2 dargestellt. Es handelt sich um ein Informationssystem, bei dem die Marketing-Manager Outputs erhalten, die sie in die Lage versetzen, Entscheidungen und Maßnahmen in bezug auf den Absatzmarkt zu treffen. Die Komponenten des Marketing-Informationssystems, und zwar die Ein-/Ausgabeeinheiten, die Methoden-, Modell- und Datenbank sind durch Input-Output-Beziehungen miteinander verbunden. Inputs können von Marketing-Managern mittels Eingabeeinheiten unmittelbar in die Datenbank eingegeben und bei Bedarf abgerufen werden. Da die Datenbank mit der Methoden- und Modellbank in Input-Output-Relation steht, verfügen die Marketing-Manager über die Möglichkeit, die jeweils in Methoden und Modellen transformierten Inputs aus der Datenbank über ihre Ausgabeeinheit abzufragen. Darüber hinaus ist es mittels der in der Methodenbank enthaltenen Methoden möglich, Inputs so aufzubereiten, daß sie dann in Marketing-Modellen Anwendung finden können. Weiterhin geht die Abbildung 2 davon aus, daß in der Modellbank Modelle vorhanden sind, die Routineentscheidungen ohne Zutun der Marketing-Manager ermöglichen.

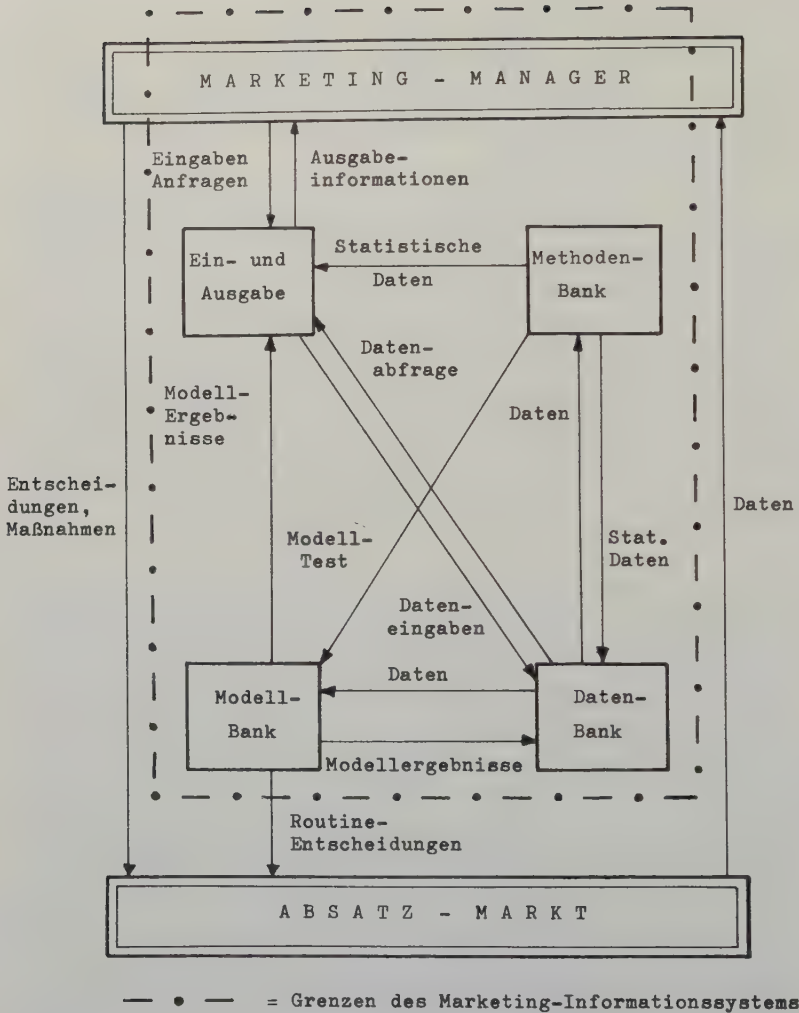


Abb. 2: Struktur und Elemente eines Marketing-Informationssystems

1 Quelle: Montgomery, D.B./Urban, G.L., Marketing Decision-Information Systems: An Emerging View, in: Journal of Marketing Research, Vol. VII, 1970, S. 226ff., hier S. 227. (Die Abbildung wurde vom Verf. geringfügig verändert.)

Von einem derartigen Marketing-Informationssystem wird im folgenden ausgegangen. Es ergibt sich aufgrund der vorgenommenen Strukturierung die Frage, ob es nicht spezielle Forschungsansätze gibt, die sich schwergewichtig mit bestimmten Komponenten dieses Systems befassen.

2. Forschungsansätze bei Marketing-Informationssystemen

In Übereinstimmung mit den bedeutendsten Definiens eines Marketing-Informationssystems, und zwar der Marketing-zweck-Orientierung, insbesondere beim Einsatz von statistischen Methoden und Marketing-Modellen, der datenverarbeitungsspezifischen Strukturierung und der Beachtung von kommunikativen Beziehungen in der Marketing-Organisation, lassen sich bei Marketing-Informationssystemen drei Forschungsansätze unterscheiden.¹ Es handelt sich dabei um den managementorientierten, den datenverarbeitungsorientierten und den systemorientierten Ansatz. Diesen drei Ansätzen kommen im Rahmen von Marketing-Informationssystemen jeweils spezifische Aufgaben zu. Diese Aufgaben gilt es im folgenden zu analysieren.

Ein Ziel von Marketing-Informationssystemen ist die Versorgung von Entscheidungsträgern mit entscheidungsrelevanten Informationen. Frage ist, ob und inwieweit die drei Forschungsansätze an der Erreichung dieses Zieles jeweils beteiligt sind. Darüber hinaus ist sowohl bei dem managementorientierten als auch bei dem datenverarbeitungsorientierten Ansatz die Verwendung eines spezifischen Datenbegriffs typisch. Dieser jeweilige Datenbegriff soll zur Charakterisierung dieser beiden Forschungsansätze näher erläutert werden.

Der managementorientierte Forschungsansatz geht von der Überlegung aus, daß ein Marketing-Informationssystem primär die Informationsbedürfnisse des Marketing-Managements zu befriedigen

¹ Vgl. zu diesen Ansätzen Blumenthal, S.C., Management Information Systems - A Framework for Planning and Development, Englewood Cliffs, N.J. 1969, S. 30; Grochla, E., Die Integration der Datenverarbeitung, in: Bürotechnik und Automation, 9. Jg., 1968, S. 108ff, hier S. 120; Young, S., Organization as a Total System, in: Cleland, D.I. / King, W.R., Hrgg., Systems-Organizations-Analysis-Management, New York usw. 1969, S. 51ff, hier S. 52f.

hat. Der Entscheidungsprozeßbezug steht somit bei diesem Ansatz im Vordergrund. Der von ihm verwendete Datenbegriff wird ebenfalls entscheidungsprozeßbezogen gesehen. Auf diesen Aspekt weist insbesondere HEINEN¹ hin, der Daten als Begrenzungsfaktoren unternehmerischen Handelns interpretiert, die zum Ausschluß bestimmter Alternativen und damit zur Eingengung des Entscheidungsspielraums führen. Es geht hier in erster Linie um die zu vollziehenden Wahlakte, die unter Beachtung des Dateneinflusses im Rahmen einer zielorientierten Marketing-Politik Bedeutung für zu fällende Entscheidungen haben. An diese Überlegung knüpft der managementorientierte Ansatz an. Für ihn ist eine Datenkonstellation unter dem Aspekt der Entscheidungsrelevanz zu kategorisieren. Diese Kategorisierung erfolgt dabei nach der jeweiligen Auswirkung auf die Entscheidungssituation im Marketing. Es ergibt sich dabei die Konsequenz, daß die für eine bestimmte Entscheidungssituation in Frage kommende Datenkonstellation nicht auf anders gelagerte Entscheidungssituationen übertragbar ist.² Eine approximative Lösung dieser Problematik ist nur dann möglich, wenn Daten so beschafft werden, daß aufgrund eines hohen Datenvolumens zu erwarten ist, daß bestimmte Datenkategorien bei alternativen Entscheidungsprozessen Anwendung finden können.

Aufgrund des Bezugs zum Entscheidungsprozeß stehen beim managementorientierten Ansatz Probleme der Informationsbedarfsanalyse im Vordergrund. Die Informationsbedarfsermittlung wird unter dem Aspekt der Erhebungsmethoden wie Befragungen, Arbeitsplatzbeschreibungen und Vermaschung beider Verfahren diskutiert.³ Besondere Beachtung findet die organisationale

1 Vgl. Heinen, E., Das Zielsystem der Unternehmung - Grundlagen betriebswirtschaftlicher Entscheidungen, Wiesbaden 1966, S. 22 und 32.

2 Vgl. Loitsberger, E., Zum Informationsbegriff und zur Frage der Auswahlkriterien von Informationsprozessen, in: Loitsberger, E., Hrg., Empirische Betriebswirtschaftslehre, Festschrift zum 60. Geburtstag von L.L. Illetschko, Wiesbaden 1963, S. 115ff, hier S. 117.

3 Vgl. z.B. Enger, N.L., Putting MIS to Work - Managing the Management Information System, New York 1969, S. 45.

Lokalisierung des Entscheidungsträgers in der Organisationspyramide. Darüber hinaus werden aggregationstheoretische Überlegungen bei den einem Entscheidungsträger zugehenden Daten, dynamische Aspekte der Informationsbedarfsermittlung und allgemeine statistisch-methodologische Probleme erörtert.¹

An die Erfassung des Informationsbedarfs schließt sich die Analyse von Informationsquellen an, aus denen der Informationsbedarf befriedigt werden kann. Aufgrund der Vielzahl möglicher Klassifizierungen von Informationsquellen und -arten wird hier oft mittels der Gegensatzpaar-Systematik versucht, eine widerspruchsfreie Klassifizierung zu erreichen.

Der letzte vom managementorientierten Ansatz oftmals vorgeschlagene Schritt besteht in der Erstellung eines Datenverzeichnisses, das sämtliche den Informationsempfängern aus den ermittelten Informationsquellen zugehenden Daten enthält. Darüber hinaus findet in der Regel eine Informationsbewertung² mittels Dringlichkeitsskalen statt, die darüber Auskunft gibt, welche Entscheidungen - und damit welche Informationsbeschaffungsaktivitäten - Priorität besitzen.

Aufgrund der Tatsache, daß die für eine bestimmte Entscheidungssituation in Frage kommenden Daten nicht auf anders gelagerte Entscheidungssituationen übertragbar sind, steht der managementorientierte Ansatz vor dem Dilemma, daß er die Relevanz von zu beschaffenden Daten a-priori nicht beurteilen kann. Eine Beurteilung der Relevanz würde voraussetzen, daß bereits zum Beschaffungs-Zeitpunkt vollständige

1 Vgl. dazu Wahl, M.P., Grundlagen eines Management-Informationssystems: Einige wesentliche Definitionen, Sonderdruck aus IBM Nachrichten, 19. Jg., 1969, S. 5.

2 Vgl. dazu Hirsch, R.E., Informationswert und -kosten und deren Beeinflussung, in: Zeitschrift für betriebswirtschaftliche Forschung, 20. Jg., 1968, S. 670ff.

Information über die zu treffenden Entscheidungen und die dafür notwendigen Daten besteht.

Angesichts der Tatsache, daß in empirischen Situationen mit beschafften Daten und agierenden Entscheidungsträgern nur Vermutungen über die Relevanz von Daten angestellt werden können, ergibt sich die Schlußfolgerung, daß in Situationen ohne empirische Daten und ohne sich vollziehende Entscheidungsprozesse das Relevanzproblem erst recht unlösbar sein wird.

Diesem Dilemma wird vom managementorientierten Ansatz dadurch begegnet, daß im Rahmen der Methodenbank eines Marketing-Informationssystems auch solche Methoden zur Anwendung gelangen sollen, die Urteils-, und damit für den Entscheidungsträger relevante Daten, verarbeiten können. "Dieses subjektive Urteil sollte daher in jedem Entscheidungskalkül zur Basis der Entscheidung gemacht werden."¹ Darüber hinaus wird vom managementorientierten Ansatz bezüglich der Marketing-Modelle, die zur Anwendung gelangen sollen, die Forderung nach bestimmten Kriterien erhoben. Dieser unter dem Begriff "Decision Calculus"² firmierende Ansatz soll die Marketing-Manager in die Lage versetzen, benutzerorientierte Modelle zu selektieren und zu entwickeln. Es soll somit gewährleistet werden, daß das Relevanzproblem wenigstens approximativ gelöst wird.

Beim datenverarbeitungsorientierten Forschungsansatz dominieren datenorganisatorische und -verarbeitungstechnische Überlegungen. Er befaßt sich primär mit der datenverarbeitungsspezifischen Organisation der Datenbank des Marketing-Informationssystems.

¹ Krautter, J., Marketing-Entscheidungsmodelle, Bd. II der "Schriften zur theoretischen und angewandten Betriebswirtschaftslehre", Hrsg. L. Pack, Wiesbaden 1973, S. 97.

² Vgl. dazu den Punkt V./2.2. dieser Arbeit.

Der hier verwendete Datenbegriff ist der der elektronischen Datenverarbeitung¹, der ohne Berücksichtigung eines entscheidungsprozeßbezogenen Sinngehaltes verwendet wird.² Im Sprachgebrauch der elektronischen Datenverarbeitung liegt ein Datum dann vor, "wenn Zuordnungskriterien in einem .. Datenelement mit Werten zusammengefaßt sind."³ Aus diesem Grund werden die Begriffe Daten und Informationen vielfach inhaltlich gleichgesetzt.

Da Daten im Rahmen der Datenverarbeitung jedoch erst dann einen "Wert" haben, wenn sie nach bestimmten Regeln verarbeitet werden können, stehen beim datenverarbeitungsorientierten Ansatz diese Verarbeitungsregeln im Vordergrund des Interesses. Problem ist, daß die Daten gespeichert, bei Bedarf wiederaufgefunden und dem Benutzer zugänglich gemacht werden müssen. Aufgrund dieser Problemstellung interessieren hier insbesondere effiziente Speicherungs- und Zugriffsarten. Eine Benutzerorientiertheit eines Marketing-Informationssystems ist dann erreicht, wenn es gelingt, benutzerorientierte Abfragesprachen und Datenverwaltungsprogramme zu entwickeln, die es dem Benutzer ermöglichen, ohne spezifische Datenverarbeitungskenntnisse Daten aus der Datenbank zu extrahieren, neue Daten aufzunehmen und irrelevante Daten zu eliminieren.

Den Bezugsrahmen für eine methodologisch orientierte Diskussion von Problemen eines Marketing-Informationssystems,

1 Eine ausführliche Darstellung des Datenbegriffs im Rahmen der Datenverarbeitung findet sich bei Schweiker, K.F., Grundlagen einer Theorie betrieblicher Datenverarbeitung, Bd. 2 der Schriftenreihe des Betriebswirtschaftlichen Instituts für Organisation und Automation an der Universität zu Köln, Hrg. E. Grochla, Wiesbaden 1966, S. 17ff.

2 Vgl. hierzu Will, H.J., Betriebswirtschaftliche Informationssystem - Versuch einer intelligenztechnischen Definition, in: Zeitschrift für betriebswirtschaftliche Forschung, 20. Jg., 1968, S. 648ff, hier S. 649.

3 Lutz, Th. / Klimech, H., Die Datenbank im Informationssystem, München-Wien 1971, S. 25.

insbesondere im Zusammenhang mit dessen Objektsystem - dem Absatzmarkt - liefert der *systemorientierte Forschungsansatz*. Ihm kommt als deskriptivem Bezugsrahmen eine wichtige heuristische Funktion zu.¹ Darüber hinaus leistet er Beschreibungs- und Systematisierungsfunktionen bei inner- und außerbetrieblichen Kommunikationsbeziehungen, d.h. bei den Kommunikationskanaltypen², über die die Entscheidungsträger miteinander kommunizieren.

Es ist mittels des systemorientierten Forschungsansatzes möglich, ein Modell des Absatzmarktes und der unternehmerischen Marketing-Entscheidungen bei der Entwicklung eines Marketing-Informationssystems zu Grunde zu legen. Ausgangspunkt dieses Ansatzes ist dabei die zweckorientierte Abgrenzung des Marketing-Systems, um seine Elemente und seine Struktur analysieren zu können. Aufgrund dieser so gewonnenen Modellvorstellungen ist es möglich, die untersuchungsrelevanten Elemente und Relationen des Marketing-Systems in ihrer gegenseitigen Interdependenz zu erfassen und offenzulegen.³

Zusammenfassend lassen sich die drei skizzierten Forschungsansätze wie folgt charakterisieren: Trotz im Detail unterschiedlicher Intentionen - diese kommen dadurch zum Ausdruck, daß innerhalb eines bestimmten Ansatzes spezielle Aspekte schwerpunktartig untersucht werden, während Aspek-

1 Vgl. Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, in: Wissenschaftsprogramm und Ausbildungsziele der Betriebswirtschaftslehre, Tagungsberichte des Verbandes der Hochschullehrer für Betriebswirtschaft e.V., Bd. I, hrsg. von G.v. Kortzfleisch, Berlin 1971, S. 167ff, hier S.168.

2 Vgl. zu den Kommunikationskanaltypen z.B. Simon, H.A., Administrative Behavior - A Study of Decision-Making Processes in Administrative Organization, 2. Aufl., New York 1957, S. 155ff.

3 Vgl. Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, a.a.O., S. 168.

te der übrigen Ansätze nur peripher von Interesse sind oder vernachlässigt werden - ist den Forschungsansätzen gemein, daß sie auf recht unterschiedliche Weise die Benutzerorientiertheit des Marketing-Informationssystem gewährleisten wollen. Aufgrund dieses gemeinsamen Ziels stehen die Forschungsansätze nicht isoliert nebeneinander.

3. Problemstellung und Gang der Untersuchung

Die einem Marketing-Informationssystem zugrundeliegende Komplexität läßt sich dann reduzieren, wenn sich sowohl der managementorientierte, der datenverarbeitungsorientierte als auch der systemorientierte Ansatz auf bestimmte Problemkreise beschränken, für deren Lösung sie jeweils geeignet sind. Mit der Behandlung der drei Forschungsansätze sind zwangsläufig verschiedene Betrachtungsebenen festgelegt. Beim management- und systemorientierten Ansatz bewegt sich die Diskussion auf pragmatisch-informationeller Ebene, während beim datenverarbeitungsorientierten Ansatz syntaktisch-informationelle Aspekte im Vordergrund stehen.

Die mit den Forschungsansätzen bereits aufgezeigten Problemschichten bilden den Gegenstand der vorliegenden Arbeit. Ziel der Arbeit ist es, einen methodologischen Beitrag zur Analyse von Datenverarbeitungsprozessen in Marketing-Informationssystemen zu erbringen, der es ermöglicht, zu einer verstärkten theoretischen Durchdringung und Strukturierung dieser Prozesse zu gelangen. Unter Berücksichtigung spezifischer Forschungsergebnisse bei den drei Ansätzen sollen insbesondere die Entscheidungsträger im Marketing in die Lage versetzt werden, die entscheidungsverbessernden Möglichkeiten und Grenzen eines Marketing-Informationssystems zu erkennen. Darüber hinaus wird in

der vorliegenden Arbeit beabsichtigt, in einer gekoppelten Analyse von Daten-, Methoden- und Modellbankproblemen einen Komplex in seiner Gesamtheit zu untersuchen, der in der Literatur bislang primär mittels partieller Analysen behandelt wurde.

Die Zielsetzung dieser Arbeit macht es erforderlich, einen formalen Bezugsrahmen zu entwickeln, Dabei wird vom systemorientierten Ansatz ausgegangen. Innerhalb dieses Bezugsrahmens können insbesondere die Probleme des Inputs in ein Marketing-Informationssystem mit der notwendigen Allgemeingültigkeit behandelt werden. Die informationsgewinnenden und -verarbeitenden Aktivitäten werden dabei als Objektentscheidungsprozesse¹ aufgefaßt.

Objektsystem des Marketing ist das Marketing-Transaktionssystem², das sich in das Marketing-Managementsystem und das Absatzmarktsystem zerlegen läßt. Sowohl die innere Struktur des Absatzmarktsystems als auch die intraorganisationalen Objektentscheidungsprozesse im Marketing-Managementsystem bleiben im Rahmen der Systembetrachtung weitgehend unberücksichtigt. Während im Marketing-Managementsystem die Objektentscheidungsprozesse in seinen Subsystemen im Vordergrund stehen, dominieren beim Absatzmarktsystem die heuristischen Funktionen der Systemanalyse. Diese gestatten es, Art, Anzahl und Interdependenzen zwischen den Systemelementen offenzulegen.³ Daran anknüpfend ergibt sich für die Datenbeschaffung die Konsequenz, daß sie sich auf die durch die Systemabgrenzung geschaffene Ordnung stützen kann. Der Problemkreis der Datenbeschaffung leitet

1 Vgl. Kirsch, W. / Meffert, H., Organisationstheorien und Betriebswirtschaftslehre, Bd. 1 der Schriftenreihe der Zeitschrift für Betriebswirtschaft, Hrg. E. Gutenberg, Wiesbaden 1970, S. 41.

2 Vgl. dazu Meffert, H., Computergestützte Marketing-Informationssysteme und Marketing-Modelle: Gegenwärtiger Stand und Entwicklungstendenzen, unveröffentlichtes Manuskript eines Beitrages zum Wirtschaftsinformatiksymposium 1973 der IBM Deutschland über computergestützte Marketing-Planung.

3 Vgl. derselbe, Systemtheorie aus betriebswirtschaftlicher Sicht, a.a.O., S. 180ff.

zu den wichtigsten Anforderungen an die zu beschaffenden Inputdaten eines Marketing-Informationssystems über.

Im weiteren sind die bislang als "Black Box" behandelten Komponenten eines Marketing-Informationssystems zu "öffnen", um die mit der Daten-, Methoden- und Modellbank verbundenen Probleme aufzudecken. Im Rahmen des datenverarbeitungsorientierten Ansatzes sind bei der Datenbank zunächst Probleme der Datenorganisation auf syntaktisch-informationeller Ebene zu lösen. Aussagen über den Inhalt der Datenbank im pragmatischen Sinn sollen vernachlässigt werden. Vielmehr bestehen Probleme auf Datenbankebene darin¹, wie die in ihr enthaltenen Daten zweckmäßig zu organisieren sind, damit dem Benutzer zweckorientierte Zugriffsmöglichkeiten geschaffen werden. Als Bezugsrahmen und Ausgangspunkt der Diskussion von Datenorganisations-Problemen erweisen sich Erkenntnisse der Sprachanalyse als geeignet. Ohne bei diesem Bezugsrahmen auf datenverarbeitungsorientierte Gesichtspunkte einzugehen, ist es mittels sprachanalytischen Ergebnissen möglich, datenverarbeitungsorientierte Begriffe wie "Verknüpfung" und "Verkettung" darzulegen. Darüber hinaus haben diese sprachanalytischen Überlegungen Konsequenzen für die zentrale Frage der Datenverarbeitung, inwieweit die logische Strukturierung von Daten mit ihrer Speicherungsform auf Datenträgern übereinstimmt und wann aus welchen Gründen davon abgewichen wird.

Weitere Probleme, die im Rahmen der Datenbank zu erörtern sind, betreffen deren Verwaltung. Es geht hier vor allem um die Fragestellung, wie die Benutzer mit der Datenbank kommunizieren und an sie Direktanfragen stellen können.

1 "Das Problem der Datenbank als eines Instruments und selbständigen Teils des Informationssystems ist keineswegs sein möglicher Inhalt". Lutz, Th., Formale Aspekte des Informationssystems und Schlußfolgerungen für ein MIS, in: IBM Nachrichten, 22 Jg., 1972, S. 54ff, hier S. 58.

Weiterhin sind im Rahmen der Datenbankverwaltung spezifische Probleme der Datenbankanpassung (Fortschreibung und Eliminierung von Daten) und des Datenschutzes zu erörtern. Unter dem Gesichtspunkt der Datenbank-Kommunikation werden Fragen der Ein- und Ausgabeeinheiten nicht eigenständig als gesonderte Komponente eines Marketing-Informationssystems untersucht. Fragen der Kommunikationseinrichtungen sind vielmehr unter dem Primat der Datenbank mit deren übrigen Problemkreisen gemeinsam zu erörtern. Um dem Aspekt der Benutzerorientierung Rechnung zu tragen, sind schließlich Probleme der Datenbank-Software offenzulegen, die am Beispiel des Integrated Data Store verdeutlicht werden sollen.

Um insbesondere das Marketing-Management mit relevanten Informationen zu versorgen, bedarf es geeigneter Methoden zur Verdichtung der in der Datenbank vorhandenen Einzeldaten. In diesem Zusammenhang erweisen sich unter dem Postulat des managementorientierten Ansatzes Methoden als notwendig, die es im Rahmen der Methodenbank gestatten, Urteilsdaten auszuwerten, daß Interpretationsmöglichkeiten dieser Daten gegeben sind.¹ Diesem Gesichtspunkt kann insbesondere durch den Einsatz multivariater Datenanalysemethoden Rechnung getragen werden, die in Form von Computerprogrammen in einer Methodenbank zusammengefaßt sind. Nach einem Überblick und der Einordnung dieser statistischen Verfahren werden diese anhand beispielhaft angenommener Marketing-Fragestellungen dargestellt.

Ein Marketing-Informationssystem wäre aber dann "rudimentär" entwickelt, wenn es nicht gleichzeitig die Vorteile der hohen Rechengeschwindigkeit des Computers auch in Zusammenhang mit den Fortschritten bei der Entwicklung von Marketing-Modellen nutzen würde. Erst durch die Symbiose

¹ Tarski spricht in diesem Zusammenhang davon, daß der Begriff "Interpretation" häufig zur Bezeichnung des Übergangs von einem formalen (syntaktischen) zu einem semantischen System verwendet wird. Vgl. Tarski, A., *Undecidable Theories*, Amsterdam 1953, S. 21.

von gespeicherten Daten, programmierten Methoden, die eine Auswertung dieser Daten ermöglichen, und der hohen Rechengeschwindigkeit des Computers sind echte Fortschritte bei zu treffenden Marketing-Entscheidungen zu erwarten.

Um Modelle in Form von Computerprogrammen im Rahmen der Modellbank eines Marketing-Informationssystems einsetzen zu können, tritt die Frage auf, welche Anforderungen an zu selektierende und/oder zu entwickelnde Modelle zu stellen sind. Zentraler Gesichtspunkt bei der Diskussion von Selektions- und Entwicklungskriterien mathematischer Modelle ist die Konfrontation zweier Ansätze, und zwar des theoretisch-intersubjektiven und des benutzerorientierten-subjektiven Ansatzes (Decision Calculus). Während sich der erste Ansatz primär mit Homomorphieproblemen von Modellen befaßt, dominieren beim zweiten die Belange der Benutzerorientiertheit. Es gilt hier, sich kritisch mit diesen beiden Ansätzen auseinanderzusetzen und sie gegeneinander abzuwägen. Aus der Synthese beider Ansätze sind dann Schlußfolgerungen für die Modellselektion und -entwicklung abzuleiten.

Abschließend werden im Rahmen dieser Arbeit - ausgehend vom Decision Calculus - einige strategische Forschungsrichtungen bei Informationssystemen aufzuzeigen sein, die aller Wahrscheinlichkeit nach zukünftige Forschungen auf dem Gebiet der Marketing-Informationssysteme entscheidend bestimmen werden.

II. Die Datenbank im Marketing-Informationssystem

1. Datenbankbegriff und Datenbanktypen

"Unter [einer] Datenbank wird der Teil eines Informationssystems verstanden, in dem mit Hilfe der elektronischen Datenverarbeitung ausgewählte Daten zusammengeführt und in besonderen Speichern für schnelle Auskünfte, Auswertungen und Rechenoperationen verfügbar gemacht werden."¹

Anforderungen, die man an ein Informationssystem stellt, lassen sich in fast allen Fällen nur über Zusammenführung und Zentralisierung der in einem abgeschlossenen Organisationsbereich anfallenden Daten verwirklichen. Es ist notwendig, daß man "ohne systembedingte Behinderung zu jedem gespeicherten Datenelement (Satz, Segment) direkt zugreifen kann und daß man solche Elemente beliebig über Hauptordnungskriterien hinweg, die die Speicherform der Bestände festlegen, aus unterschiedlichen Dateien in Verarbeitungsprozessen kombinieren kann."²

Bei den Datenbanken lassen sich aufgrund der in ihnen enthaltenen Daten formatierte und nicht-formatierte Datenbanken unterscheiden. Der Unterschied zwischen den beiden Datenbanktypen besteht aufgrund des Formats der in ihnen enthaltenen Daten, die zu Datenfeldern zusammengefaßt werden, wobei die Größe und die Struktur dieser Datenfelder das Format bestimmen. "Es handelt sich .. um Felder, wie Werte, Nummern, Bezeichnungen, Namen und so fort, die in einem fest vordefinierten Schema angeordnet sind."³ Darüber hinaus zeichnen sich formatierte Daten dadurch aus, daß sie entsprechend den syntaktischen Regeln einer Pro-

1 Hosse, H., Landesdatenbank Nordrhein-Westfalen, in: ADL-Nachrichten, 13. Jg., 1968, S. 320ff, hier S. 320.

2 Lutz, Th., Adressketten und Datentechnik, in: Bürotechnik und Automation, 9. Jg., 1968, S. 178ff, hier S. 178.

3 Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 74.

grammiersprache transformiert werden.¹

Nicht-formatierte Daten finden sich insbesondere bei Textwiederholungssystemen (z.B. im Rahmen der Werbung). Bei diesen bestehen die Daten aus verbal dargestellten Informationen, die kein besonderes Format haben, sondern wie etwa Medienbelegungen², Patente, Nachrichten und Lebensläufe durch im Text auftretende Schlüsselwörter identifiziert werden. Im Gegensatz dazu enthalten die formatierten Datenbanken solche Daten, auf die unmittelbar aufgrund der in ihnen residierenden Ordnungskriterien zugegriffen werden kann. Konzeptionell stehen sich somit formatierte und nicht-formatierte Datenbanken konträr gegenüber. Während die ersten in den Bereich der kommerziellen Datenverarbeitung einzuordnen sind, lassen sich nicht-formatierte Datenbanken mehr in den Bereich des "Dokumentenretrieval" einordnen. Im Rahmen dieser Arbeit werden ausschließlich Probleme formatierter Datenbanken behandelt, da die meisten sich im Marketing ergebenden Daten formatiert sind.

1 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 74.

2 Als Beispiele können hier die Leserschaftsanalysen großer Publikumszeitschriften angeführt werden, bei denen eine Identifizierung von zueinandergehörigen Daten nur mittels im Text enthaltener Schlüsselworte erfolgen kann.

2. Organisation der Datenbank

2.1 Formale Datenorganisation

2.1.1 Sprachanalytische Segmentierung

Im Rahmen der Datenorganisation, deren Ziel die Schaffung von Zugriffsmöglichkeiten zu den Daten einer Datenbank ist, lassen sich zwei Ebenen unterscheiden, und zwar die logisch-sprachliche Ebene (formale Datenorganisation) und die Zeichenebene (physische Datenorganisation). Die Sprachebene umfaßt dabei die Elemente einer Sprache (Begriffe) und die Verknüpfung dieser Elemente zu logischen Einheiten (Sätzen) und die Regeln für deren Handhabung.

Auf der Zeichenebene bestehen verschiedene Möglichkeiten, sprachliche Elemente und logische Einheiten mit verschiedenen Symbolen zu bezeichnen. Elemente, logische Einheiten und Verknüpfungen auf der Sprachebene sind auf der Zeichenebene als spezifisch wahrnehmbare und in Zeichen transformierte Größen anzusehen, die aufgrund vorgegebener Regeln der Zeichenebene verarbeitet werden können. Um die Sprachebene näher zu präzisieren und um dann den Zusammenhang zwischen Sprach- und Zeichenebene aufzeigen zu können, sind neuere Ansätze in der Sprachtheorie geeignet.¹

Gegenstand der Sprachtheorie ist die Analyse von Ordnungen in natürlichen und künstlichen Sprachen², da sich nur mittels einer geordneten Sprache Aussagen über reale Objekte gewinnen lassen.³ Der Zweck, der hier mit der Sprachanalyse verfolgt wird, ist dualistisch gelagert. Einerseits soll

1 Vgl. dazu insbesondere Kamlah, W. / Lorenzen, P., Logische Propädeutik oder Vorschule des vernünftigen Redens, Mannheim 1967 und Whorf, B.L., Sprache - Denken - Wirklichkeit: Beiträge zur Metalinguistik und Sprachphilosophie, hrsg. v. P. Krausser, Reinbek 1963.

2 Vgl. Seiffert, H., Einführung in die Wissenschaftstheorie - Sprachanalyse - Deduktion - Induktion in Natur- und Sozialwissenschaften, 3. Aufl., München 1971, S. 3ff.

3 Vgl. Kamlah, W. / Lorenzen, P., Logische Propädeutik, a.a.O., S. 46f.

gezeigt werden, daß logische Gemeinsamkeiten zwischen sprachlicher Ordnung und formaler Datenorganisation bestehen. Andererseits sollen mittels der Sprachanalyse die Grundlagen für das Aufzeigen von Unterschieden zwischen der Verknüpfung von Elementen einer Sprache zu logischen Einheiten und der datenverarbeitungsorientierten Verarbeitung dieser logischen Einheiten gewonnen werden. Die Theorie der Verknüpfungen als ein Bestandteil der Sprachtheorie bietet hierfür die Möglichkeit, den theoretischen Hintergrund des skizzierten Problems näher zu durchleuchten.¹

Die Sprachanalyse arbeitet primär mit den Kategorien "Objekt", "Eigenschaft" und "Relation".² Bestimmten Objekten werden Prädikatoren (Eigenschaften und Relationen) zugeordnet, wobei dieser Vorgang als Prädikation bezeichnet wird. Als Objekt wird das definiert, was mittels eines Begriffs der Sprache bezeichnet oder dem in der Prädikation ein Prädikator zugesprochen werden kann. Je nachdem wie viele Objekte zu einem Prädikator gehören, lassen sich ein- und mehrstellige Prädikatoren unterscheiden.³ Die Beziehung zwischen mehreren Objekten wird als Relation bezeichnet. Folglich muß auch für eine Relation mit nur einem Objekt - eine einstellige Relation - eine Kennzeichnung eingeführt werden. Diese Kennzeichnung erfolgt durch den Terminus "Eigenschaft".⁴

Bei einem Vergleich der Kategorien "Objekt", "Eigenschaft" und "Relation" zeigt sich, daß die Relation als Sonderfall der Eigenschaft betrachtet werden kann. "Eigenschaft"

1 Vgl. Flechtner, H.-J., Grundbegriffe der Kybernetik - Eine Einführung, Stuttgart 1969, S. 233.

2 Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, aus dem Russischen übersetzt von G. Kröber unter Mitarbeit von P. Bollhagen, Berlin 1965, S. VII.

3 Vgl. dazu Hilbert, D. / Ackermann, W., Grundzüge der theoretischen Logik, 5. Aufl., Berlin usw. 1967, S. 141ff.

4 Vgl. Kamlah, W. / Lorenzen, P., Logische Propädeutik, a.a.O., S. 38ff.

und "Relation" sind somit analoge Begriffe, die sich ausschließlich durch die Anzahl der Objekte unterscheiden, auf die sie sich beziehen.

Die Relation genügt somit den bei der Kategorie "Eigenschaft" genannten Bedingungen. Es ist daher möglich, die Relation als eine Eigenschaft zu interpretieren, die mindestens zwei, höchstens endlich viele Objekte charakterisiert.¹ Die Tatsache, daß die Relation als Sonderfall der Eigenschaft zu gelten hat, findet auch ihre Entsprechung in der natürlichen Sprache: Die Adjektive bezeichnen Eigenschaften von Objekten; darüber hinaus bringen bestimmte Adjektive wie z.B. "gemeinschaftlich" eine Relation zwischen den Objekten zum Ausdruck.

Homogene Objekte lassen sich zu Objektmengen zusammenfassen. Objektmengen sind als geordnete Mengen von Objekten zu interpretieren. Der Prozeß der Objektidentifizierung und Objektmengenbildung ist primär untersuchungsabhängig und geschieht durch Identifizierung einzelner für die Analyse interessierender Objekte, durch Überprüfung ihrer begrifflichen Homogenität und durch anschließende Zusammenfassung der als begrifflich homogen erachteten Objekte. Schließt man beispielsweise aus der Zahl der Merkmale eines Begriffs ein oder mehrere Merkmale durch isolierende Abstraktion aus, so erhält man einen übergeordneten, fügt man den vorhandenen Merkmalen ein neues hinzu (Determination), so erhält man einen untergeordneten Begriff. Sofern ein Begriff überhaupt einer Determination fähig ist, besitzt er Allgemeingültigkeit. Diese Allgemeingültigkeit ist für die Objektmengen oder -gruppen Postulat. Das Spezifikum dieses kategorialen sprachanalytischen Begriffsinstrumentariums besteht darin, daß es äußerst universell ist. Die bislang gewonnenen Feststellungen sollen im folgenden anhand eines dem Marketing entlehnten Beispiels verdeutlicht werden.

¹ Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 54f.

Im Rahmen einer sozio-ökonomischen Marktsegmentierung geht es um das Problem, eine Objektmenge "Personen" durch Eigenschaften zu charakterisieren, die für die segmentierende Unternehmung von Interesse sind. Diese Eigenschaften lassen sich zu einer Charakteristik zusammenfassen, worunter die Aggregation homogener charakterisierender Elemente (Eigenschaften) verstanden werden soll. Jede Objektmenge oder -gruppe läßt sich im Rahmen einer sozio-ökonomischen Marktsegmentierung durch eine quasi unbeschränkte Anzahl verschiedener Eigenschaften charakterisieren. Aus Operationalitätsgründen beschränkt sich die Analyse jedoch auf eine begrenzte Anzahl charakterisierender Eigenschaften¹, so daß jede Objektgruppe durch eine Charakteristik möglichst exakt beschrieben werden kann.

Die Charakteristiken der Objekte lassen sich in ein- und multidimensionale, in quantitative und qualitative und in abhängige und unabhängige einteilen.² Eine eindimensionale Charakteristik ist dann vorhanden, wenn nur ein charakterisierendes Element vorhanden ist.³ Zwei Charakteristiken sind abhängig, wenn irgendein Zustand eines der Charakteristiken die Anzahl möglicher Zustände anderer Charakteristiken beschränkt. Dieses soll anhand eines Beispiels näher verdeutlicht werden:

-
- 1 "Jede Person besitzt eine Anzahl von Charakteristika, die für sie spezifisch sind; Wenn wir unter P eine Eigenschaft verstehen, die nur einer Person zukommt, dann können wir sagen: 'Ich gebe den Namen "A" der Person, die die Eigenschaft P besitzt'." Russell, B., Human Knowledge, London 1951, S. 318f.
 - 2 Zu den Charakteristiken bzw. Eigenschaftstypen vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 98ff und S. 103ff.
 - 3 Ein- und multidimensionale Charakteristiken werden im folgenden durch Vektoren beschrieben. In der Mathematik werden n-dimensionale Eigenschaften durch Tensoren ausgedrückt. Tensoren 0-ter Stufe sind Skalare, Tensoren 1-ter Stufe Vektoren. Tensoren jeder Stufe sind Operatoren, so daß sich ein Tensor zweiter Stufe als Matrix auffassen läßt. Vgl. dazu Teichmann, H., Physikalische Anwendung der Vektor- und Tensorrechnung, Mannheim 1964, S. 31ff.

Im Rahmen einer sozio-ökonomischen Marktsegmentierung eines Automobilherstellers ist in erster Linie nur der Teil der Bevölkerung interessant, der einen Führerschein besitzt. Das Besitzen eines Führerscheins ist aber von einem bestimmten Mindestalter abhängig. Im folgenden soll C_{11} die Charakteristik "besitzt einen Führerschein" und C_{21} die Charakteristik "Alter" sein. Die erste Charakteristik besitzt die charakterisierenden Eigenschaften S_{11} "besitzt einen Führerschein" und S_{12} "besitzt keinen Führerschein". Die zweite Charakteristik umfaßt die Eigenschaften S_{21} "Alter unter 18 Jahren" und S_{22} "Alter über 18 Jahren".

Jede Charakteristik setzt sich aus Charakterisierungsvektoren zusammen.¹ Ein Charakteristikvektor besteht aus einer Menge geordneter Charakteristikelemente qualitativer und quantitativer Art.² Die Aufgabe der Charakterisierungsvektoren besteht in einer möglichst präzisen Charakterisierung der Objektgruppen. Erst durch den Charakterisierungsvektor läßt sich der Begriff der Charakteristik eindeutig fassen: Eine Charakteristik ist die Zusammenfassung qualitativer und quantitativer Charakteristikvektoren. Da die Anzahl der in die Analyse einbeziehbaren Objektgruppen als endlich angenommen wird, lassen sich diese durch eine endliche Anzahl von Charakteristikvektoren beschreiben. Dabei ist es möglich, daß sich die Charakteristiken einer Objektgruppe überschneiden können. Aus diesem Grund ist für eine überschneidungsfreie Charakteristik - so auch bei der sozio-ökonomischen Marktsegmentierung - zu sorgen.

Die Charakteristiken lassen sich durch Vektoren bzw. Operatoren ausdrücken, wobei die Vektorelemente S_{ij} den Zustand j für die Charakteristik i angeben. Dieses soll anhand des Ausgangsbeispiels näher verdeutlicht werden. Es ergeben sich folgende primäre Charakteristikvektoren:

¹ Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 120 und S. 124.

² Vgl. hierzu S. 32 dieser Arbeit.

$$C_{11} = \begin{vmatrix} S_{11} \\ S_{12} \end{vmatrix} \quad \begin{array}{l} \text{(besitzt einen Führerschein)} \\ \text{(besitzt keinen Führerschein)} \end{array}$$

und

$$C_{21} = \begin{vmatrix} S_{21} \\ S_{22} \end{vmatrix} \quad \begin{array}{l} \text{(jünger als 18 Jahre)} \\ \text{(älter als 18 Jahre)} \end{array} .$$

Nach dem Exklusionsverhältnis der Elemente eines Charakteristikvektors zu den Elementen eines anderen lassen sich abhängige und unabhängige Charakteristiken, die jeweils noch in paarweise und nicht-paarweise Abhängigkeiten/Unabhängigkeiten unterteilt werden können, unterscheiden. Eine Charakteristik ist dann abhängig, wenn ein Element eines Charakteristikvektors zu einem anderen Element eines anderen Charakteristikvektors in Exklusion steht. Die paarweise und mehrfache Abhängigkeit bezieht sich dann auf die bi- bzw. multilaterale Exklusion zweier oder mehrerer Elemente von zwei oder mehreren Charakteristiken. Dieses impliziert bei paarweiser Abhängigkeit, daß der Zustand der einen Charakteristik vom Zustand einer anderen Charakteristik abhängig ist. Es ergibt sich, daß ein Charakteristikvektor seinen Zustand (d.h. die Ausprägung seiner Charakteristikelemente) unabhängig vom Zustand des zweiten Charakteristikvektors nicht aufrechterhalten kann. Ein Beispiel für paarweise unabhängige Charakteristiken sind die Eigenschaften "Geschlecht" und "besitzt einen Führerschein".

Die paarweise Abhängigkeit bzw. Unabhängigkeit kann zwischen zwei und mehreren Charakteristikvektoren durch einen Abhängigkeitsvektor D_{ij}^{j1} , der die Abhängigkeit bzw. Unabhängigkeit des Charakteristikvektors C_{ij} vom Charakteristikvektor C_{j1} angibt, zum Ausdruck gebracht werden. In diesem

Vektor ist das jeweilige Element d_{ij} gleich 1, wenn das Element i im Charakteristikvektor C_{ij} vom Element j im zweiten Charakteristikvektor $C_{j\cdot}$ paarweise unabhängig ist. Bei paarweiser Abhängigkeit ist das Vektorelement gleich Null. Wenn die jeweiligen Elemente der Charakteristikvektoren übereinstimmen, gibt der Abhängigkeitsvektor diese Abhängigkeitsbeziehungen an.¹

Der Abhängigkeitsvektor lautet im Fall des Ausgangsbeispiels:

$$D_{21}^{11} = \begin{vmatrix} s_{11}/s_{21} \\ s_{11}/s_{22} \end{vmatrix} = \begin{vmatrix} d_{11} \\ d_{12} \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \end{vmatrix}.$$

Das Ergebnis besagt, daß, wenn man jünger als 18 Jahre ist, keinen Führerschein besitzen kann ($d_{11} = 0$), während man bei höherem Lebensalter einen Führerschein besitzen kann ($d_{12} = 1$). Die Charakteristiken sind folglich gegenseitig voneinander abhängig.

Darüber hinaus sind mehrwertige Abhängigkeiten bzw. Unabhängigkeiten zwischen den Charakteristikvektoren möglich, z.B. zwischen "Alter", "besitzt einen Führerschein" und "Geschlecht", die auf analoge Weise mittels Vektoren ausgedrückt werden können.

Nach dem Grad der Unmittelbarkeit der Charakterisierung einer Objektgruppe kann zwischen primären und sekundären Charakteristikvektoren unterschieden werden. Ein primärer

¹ Zum formalen Problemkreis vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 124ff. Neben der Abbildungsmöglichkeit mittels Vektoren bzw. Operatoren existieren noch Möglichkeiten, Abhängigkeiten und Unabhängigkeiten mittels Graphen abzubilden. Vgl. dazu Berge, C., The Theory of Graphs and its Applications, London-New York 1964 und Busacker, R.G. / Saaty, Th.L., Endliche Graphen und Netzwerke, München-Wien 1968.

Charakteristikvektor¹ ist eine Menge geordneter Elemente, die in unmittelbarer Beziehung zur charakterisierenden Objektgruppe stehen, während sekundäre Charakteristikvektoren Mengen von geordneten Elementen darstellen, bei denen diese Unmittelbarkeit nicht gegeben ist. So ist es z.B. im Rahmen sozio-ökonomischer Marktsegmentierungen möglich, verschiedene psychographische Segmentierungskriterien zu definieren, die mehr oder weniger unmittelbar das abzugrenzende Marktsegment festlegen.

Darüber hinaus kann zwischen ein- und multidimensionalen Charakteristiken unterschieden werden. Bei einer eindimensionalen Charakteristik besteht der Charakteristikvektor einer bestimmten Objektgruppe nur aus einem Element², es ist also nur ein Zustand der Objektgruppe möglich, während bei einer multidimensionalen Charakteristik mehr als nur ein Zustand durch den Charakteristikvektor angegeben werden kann.

Weiterhin ist zwischen qualitativen und quantitativen Charakteristikvektoren zu unterscheiden. Es ist relativ unproblematisch, Charakteristikvektoren für eine Charakteristik mit quantitativer Bedeutung zu formulieren, wenn die Zustände, d.h. Elemente der Vektoren, durch nominal und/ oder metrisch skalierte Größen ausgedrückt werden können.³ Bei kontinuierlicher Spannweite dieser Größen werden die Zustände als geeignete Intervalle definiert, wie z.B. bei Alters- und Einkommensklassen. Charakteristikvektoren für Charakteristiken mit qualitativer Bedeu-

1 Die Einteilung der Charakteristikvektoren erfolgt in Anlehnung an Perreault, J.M., Towards a theory for UDC - Essays aimed at structural understanding and operational improvement, London 1969, S. 119 ff, insbesondere S. 123ff.

2 In diesem Fall liegt ein Skalar vor, der durch Angabe eines Zahlenwertes vollständig bestimmt ist.

3 Vgl. dazu Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 34f.

tung können nur mittels verschiedener Codes definiert werden.¹

Im Rahmen einer sozio-ökonomischen Marktsegmentierung soll angenommen werden, daß eine Objektgruppe "Personen" existiert, die durch vier Charakteristikvektoren, und zwar C_{11} "Geschlecht", C_{21} "Familienstand", C_{31} "Altersklasse" und C_{41} "Einkommensklasse" charakterisiert werden kann:

C_{11}	=	<table> <tr><td>männlich</td></tr> <tr><td> </td></tr> <tr><td>weiblich</td></tr> </table>	männlich		weiblich	C_{21}	=	<table> <tr><td>ledig</td></tr> <tr><td>verheiratet</td></tr> <tr><td> </td></tr> <tr><td> </td></tr> <tr><td>geschieden</td></tr> <tr><td>verwitwet</td></tr> </table>	ledig	verheiratet			geschieden	verwitwet	C_{31}	=	<table> <tr><td>0<t<20</td></tr> <tr><td>20<t<40</td></tr> <tr><td> </td></tr> <tr><td> </td></tr> <tr><td>40<t<60</td></tr> <tr><td>60<t<80</td></tr> </table>	0<t<20	20<t<40			40<t<60	60<t<80
männlich																							
weiblich																							
ledig																							
verheiratet																							
geschieden																							
verwitwet																							
0<t<20																							
20<t<40																							
40<t<60																							
60<t<80																							
C_{41}	=	<table> <tr><td>0<Y<10'</td></tr> <tr><td>10'<Y<20'</td></tr> <tr><td> </td></tr> <tr><td>20'<Y<30'</td></tr> <tr><td>30'<Y<40'</td></tr> </table>	0<Y<10'	10'<Y<20'		20'<Y<30'	30'<Y<40'																
0<Y<10'																							
10'<Y<20'																							
20'<Y<30'																							
30'<Y<40'																							

Bei diesem Beispiel sind zwei Charakteristikvektoren vorhanden, die qualitativ-klassifikatorische Elemente enthalten (Geschlecht und Familienstand) und zwei Charakteristikvektoren, die quantitative Charakteristikelemente (Alters- und Einkommensklasse) umfassen.

Es können zwischen sämtlichen Objekten der Objektgruppe "Personen" Relationen bestehen (Intraobjektgruppenrelationen), ebenso wie zwischen sämtlichen Objektgruppen Rela-

¹ Vgl. z.B. die Ansätze für Dezimalklassifikationen bei de Grolier, E., A study of general categories applicable to classification and coding in documentation, Paris 1962, S. 18 - 42 and Vickery, B.C., Classification and indexing in science, London 1959, S. 186ff.

tionen existieren können (Interobjektgruppenrelationen).¹ Eine Intraobjektgruppenrelation ist z.B. dann gegeben, wenn in der Objektgruppe j "Personen" $\{P_{1j}, P_{2j}, \dots, P_{ij}\}$ P_{1j} in Relation zu P_{ij} - $P_{1j}RP_{ij}$ - steht, also eine interne Relation zwischen den Objekten der Objektgruppe vorliegt. Die Interobjektgruppenrelation liegt dann vor, wenn eine externe Relation zwischen verschiedenen Objekten der Objektgruppen j und l - z.B. $P_{1j}RP_{il}$ - besteht.²

Nach der Richtung der vorliegenden Relation lassen sich primäre und sekundäre Objekte unterscheiden. Bezieht sich z.B. das Objekt P_{1j} auf das Objekt P_{2j} in der Objektgruppe j , so ist Objekt 1 Primärobjekt, während Objekt 2 Sekundärobjekt ist. Da beide Objekte ein und derselben Objektgruppe angehören, liegt eine interne Primärobjekt-Sekundärobjekt-Relation vor. Diese Relation kann durchaus kommutativ sein, d.h. P_{2j} kann sich auf P_{1j} beziehen und umgekehrt.

Gleichermaßen gibt es externe Primärobjekt-Sekundärobjekt-Relationen, wenn die Objekte unterschiedlichen Objektgruppen angehören. Im obigen Fall der Intraobjektgruppenrelation liegt somit eine externe Primärobjekt-Sekundärobjekt-Relation vor, und zwar mit dem Primärobjekt 1 der Objektgruppe j und dem Sekundärobjekt i der Objektgruppe l .

Die zwischen einer oder mehreren Objektgruppen existierenden Relationen lassen sich mittels Beziehungs- und Verkettungsvektoren beschreiben.³ Eine Beziehung liegt dann vor, wenn Primärobjekte in einer Objektgruppe mit

1 Ujomov spricht in diesem Zusammenhang von inneren und äußeren Relationen. Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 113f und S. 159.

2 Der Begriff der Interobjektgruppenrelation wird hier mit dem der externen Relation i.S. Ujomovs gleichgesetzt. Vgl. dazu Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 159.

3 Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 127.

Sekundärobjekten derselben (interne Beziehung) oder einer anderen Objektgruppe (externe Beziehung) verbunden sind. Die Richtung der Beziehung bestimmt darüber, welche Objekte als primäre oder sekundäre anzusehen sind. Das Registrieren dieser Beziehung impliziert, daß auf jedes Primärobjekt sekundäre Objekte bezogen sind. Das heißt, daß in einer Sprache ein Prädikator (Eigenschaft oder Relation) immer auf ein ihm zugehöriges Objekt verweist. Nur wenn diese Reihenfolge "Objekt-Prädikator-Objekt" gegeben ist, sind logische Aussagen auf Sprachebene möglich. Es existiert auf der Ebene der Sprache somit nicht das Problem, daß zueinander gehörige Prädikationen in verschiedenen Sätzen stehen. Die physische Begriffsreihenfolge stimmt in einer Sprache mit dem logischen Aufbau ihrer Sätze überein.

Eine Beziehung kann immer dann umgekehrt werden, wenn die durch sie begründete Verbindung kommutativ ist.¹ So ist es z.B. möglich, in der Objektgruppe "Personen" die Beziehung zwischen Eltern und Kindern zu berücksichtigen, und zwar mit den "Eltern" als Primärobjekten und den "Kindern" als Sekundärobjekten. Ist die Beziehung zwischen "Eltern" und "Kindern" definiert, kann jede Person danach registriert werden, ob sie Kinder hat oder nicht.² In einer solchen internen Beziehung kann ein bestimmtes Objekt in einer Beziehung als Primärobjekt (Eltern) und als Sekundärobjekt (Kinder) in einer anderen Beziehung vorkommen. Die inverse Beziehung besteht zwischen den Primärobjekten "Kinder" und den Sekundärobjekten "Eltern", so daß hier eine Registrierung jedes Objekts der Objekt-

1 Zum Kommutativitätsproblem vgl. Behnke, H. et al., Mathematik, Bd. 1, Frankfurt 1964, S. 249 und S. 314f.

2 "Wenn wir von den .. [Objekten] zur Relation übergehen, dann ist die Stellenzahl der Relation schon durch die Definition der Relation genau fixiert. Wenn wir dagegen von einer selbständig definierten Relation ausgehen, kann die Zahl der .. [Objekte], zwischen denen diese Relation besteht, unbestimmt sein. "Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 115. (Im Original teilweise kursiv.)

gruppe entweder als "Eltern" oder "keine Eltern" möglich ist.

Es ist einleuchtend, daß eine Beziehung und die zugehörige inverse Beziehung von gleicher Bedeutung sind, soweit es ihre Informationsinhalte anbelangt. Darüber hinaus können die Beziehung und die inverse Beziehung auseinander abgeleitet werden. Die Wahl der "Beziehung" oder der "inversen Beziehung" ist dabei zweckabhängig und wird durch die am häufigsten gebrachte Richtung der Beziehung innerhalb der Objekte bestimmt. In diesem Sinne kann die Beziehung im Prinzip als Charakteristik interpretiert werden, und zwar als "Status", den die Objekte innerhalb der Objektgruppe als Sekundärobjekte besitzen. Der Beziehungsvektor gibt somit die Beziehungen an, die zwischen einem Objekt einer oder mehrerer Objektgruppen vorliegen.

Im Gegensatz zur Beziehung stellt die Verkettung darauf ab, daß Verknüpfungen zwischen Objekten einer oder mehreren Objektgruppen vorliegen, die dasselbe Objekt als Primärobjekt besitzen.¹ Die externe Primärobjekt-Sekundärobjekt-Relation stellt dann den Übergang zur Verkettung dar, wenn in einer Objektgruppe ein Objekt als Primärobjekt vorhanden ist, zu dem die Objekte anderer Objektgruppen in Sekundärobjekt-Relation stehen. Werden Beziehung und Verkettung zusammen betrachtet, so kommt die Verkettung durch vertikale Struktur der Objekthierarchie, die Beziehung durch die horizontale Struktur dieser Hierarchie zum Ausdruck.

Daraus folgt, daß die Beziehung in einer Sprache durch horizontale Strukturen (die Prädikation findet zwischen Objekten einer Objektgruppe statt), die Verkettung durch ver-

¹ Vgl. dazu das Gedankenexperiment bei Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 93ff und die Relationshierarchie bei Perreault, Towards a theory for UDC, a.a.O., S. 127 und S. 130f.

tionale Strukturen (die Prädikation findet zwischen Objekten verschiedener Objektgruppen statt) erfolgt.

Die möglichen Beziehungen und Verkettungen innerhalb einer Objektgruppe können durch einen Relationsvektor zum Ausdruck gebracht werden. Dieser Vektor enthält sämtliche Objekte einer Objektgruppe. Unter der Annahme, daß die Objektgruppe j "Personen" aus fünf Objekten $P_{1j}, P_{2j}, \dots, P_{5j}$ besteht, zwischen denen Relationen vorliegen können, enthält der Relationsvektor ebenfalls fünf Elemente.¹

Durch Festlegung der Reihenfolge der Charakteristikvektoren und der Zuordnung der Charakteristikelemente auf diese einzelnen Charakteristikvektoren ist es prinzipiell möglich, eine Charakterisierung für beliebig viele Objektgruppen mit beliebig vielen Charakteristikelementen vorzunehmen.² Es kann z.B. festgelegt werden, daß ein zutreffendes Charakteristikelement die 1, anderenfalls die Null zugewiesen bekommt, so daß auch bei qualitativen Charakteristikelementen eine numerische Kennzeichnung möglich ist. So ist es z.B. in der Objektgruppe "Personen" möglich, daß die Person P_1 , die Vater³ von P_4 und P_5 ist, in der Altersklasse von $20 < t < 40$ liegt und ein Einkommen in der Klasse von $20' < Y < 30'$ bezieht, numerisch zu charakterisieren. Durch Einsetzen der oben bezeichneten Symbole für ein zutreffendes (1) bzw. nichtzutreffendes (0) Charakteristikelement werden die Charakteristikvektoren in Zustandsvektoren überführt. Zustandsvektoren sollen

1 Vgl. zur binären Darstellungsmöglichkeit von Relationsmatrizen Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 124ff.

2 Hierbei ist zu beachten, daß ein Objekt, das in Relation zu anderen Objekten steht, nicht zu sich selbst in Relation stehen darf; ansonsten ist Russellsche Antinomie möglich. Vgl. dazu Behnke, H. et al., Mathematik, Bd. 1, a.a.O., S. 248.

3 Die durch den Begriff "Vater" gekennzeichnete Relation ist ein- und mehrstellig. Darüber hinaus wird angenommen, daß dieser "Vater" Ehemann ist, so daß die hier vorliegende Relation zweistellig ist. Folglich kann in diesem Fall dem zweiten Element des zweiten Charakterisierungsvektors die 1 zugewiesen werden.

als numerische Konkretisierung eines Charakteristikvektors verstanden werden. Für die Person P_1 lauten die Zustandsvektoren:

$$C_{11} = \begin{vmatrix} 1 \\ 0 \end{vmatrix} \quad C_{21} = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{vmatrix} \quad C_{31} = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{vmatrix} \quad C_{41} = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 0 \end{vmatrix} .$$

Es ist nun das Ziel, diese Zustandvektoren so umzuformulieren, daß unmittelbar auf den Zustand der Charakteristik geschlossen werden kann. Unter dem Zustand einer Charakteristik werden die für ein bestimmtes Objekt einer Objektgruppe spezifizierten Charakteristikvektoren verstanden. Durch den Zustand der Charakteristik wird gleichzeitig der Zustand des Objekts einer Objektgruppe festgelegt. Mehrere Zustände sind hierbei nicht gleichzeitig möglich¹, wohl aber können die Objekte einer Objektgruppe gleichzeitig für sich genommen unterschiedliche Zustände aufweisen.

Die Zustandsvektoren können, da sie noch redundant sind, in ein Zustandsprofil überführt werden. Das Zustandsprofil gibt die "Stelle" eines Zustandsvektor-Elements an, an der das zutreffende Charakteristikelement steht. Es erfaßt somit, welches Element in den einzelnen Zustandsvektoren für die Charakterisierung einer Objektgruppe zutrifft. Bei Kenntnis der einzelnen Zustandsvektoren genügt somit die Angabe des Zustandsprofils, um ein Objekt aus einer Objektgruppe zu charakterisieren.² Für die Person P_1

¹ Vgl. Ujomov, A.I., Dinge, Eigenschaften und Relationen, a.a.O., S. 125.

² Dieses Zustandsprofil kann noch in binäre Zahlen umgerechnet werden, so daß sämtliche zur Anwendung gelangenden Vektoren ausschließlich auf Binärcodes basieren. Zur Binärcodierung vgl. Flechtner, H.-J., Grundbegriffe der Kybernetik, a.a.O., S. 94ff.

ergibt sich folgendes Zustandsprofil:

$$Z_1 = \begin{vmatrix} 1 \\ 2 \\ 2 \\ 3 \end{vmatrix} .$$

Dieses Zustandsprofil besagt, daß das erste Charakteristikelement des Charakteristikvektors "Geschlecht", das zweite Element des Charakteristikvektors "Familienstand", das zweite Element des dritten Charakteristikvektors "Altersklasse" und das dritte Element des vierten Charakteristikvektors "Einkommensklasse" bei der Person P_1 zutreffen.

Der Gesamtzusammenhang soll im folgenden graphisch verdeutlicht werden, um die Gemeinsamkeiten zwischen Sprachebene und formaler Datenorganisation aufzuzeigen. In der Abbildung 4 werden beobachtbare Phänomene der Realität (mögliche Objekte) nach jeweiligen Zwischenschritten isolierender Abstraktion in Objekte und Charakteristikelemente eingeteilt. Bei den Charakteristikelementen können quantitative und qualitative Elemente unterschieden werden. Die Objekte werden dann entsprechend ihrer begrifflichen Homogenität in Objektgruppen zusammengefaßt. Bei diesen Objektgruppen sind Relationen zwischen Objekten einer Objektgruppe und verschiedenen Objekten unterschiedlicher Objektgruppen möglich. Die Intraobjektgruppen-Relationen können durch einen Beziehungsvektor, die Interobjektgruppen-Relationen durch einen Verkettungsvektor erfaßt werden. Die Charakteristikelemente lassen sich entsprechend ihrer Zuordnung zu den Objektgruppen zu Charakteristikvektoren zusammenfassen, wobei diese Zusammenfassung als Charakteristik der Objektgruppe bezeichnet wird. Je nachdem welche Charakteristikelemente für ein bestimmtes Objekt einer Objektgruppe zutreffen, werden durch Zuweisungen von Codes die Charakteristikvektoren in Zustands-

Charakteristik = Zusammenfassung der Charakteristikvektoren

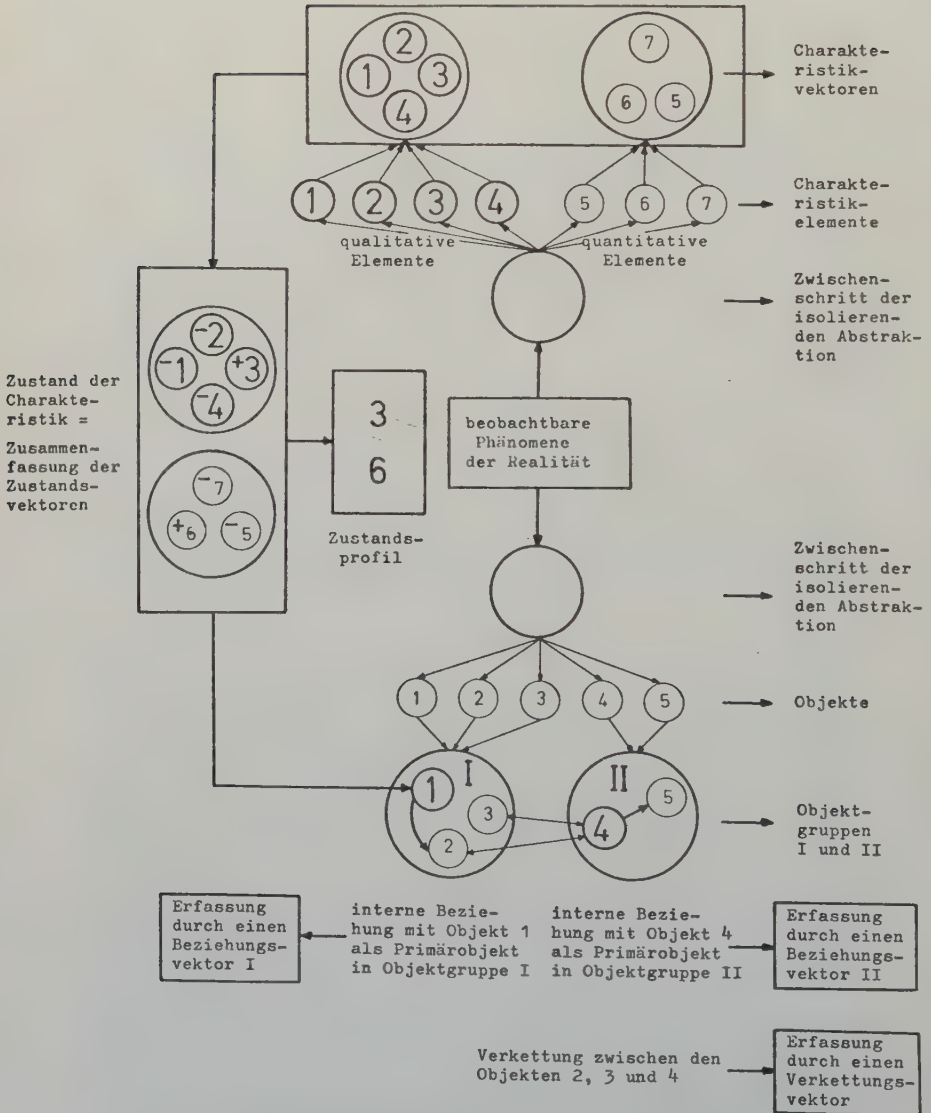


Abb. 3 : Sprachanalytische Segmentierung

vektoren überführt. Die Zusammenfassung dieser Zustandsvektoren ergibt den Zustand der Charakteristik. Der Zustand der Charakteristik kann wiederum durch die Anwendung von Codes in ein Zustandsprofil umgeformt werden, das die "Stelle" eines zutreffenden Zustandsvektor-Elements angibt.

Die Untersuchung der Sprachebene hat zu zwei bedeutsamen Feststellungen geführt. Die erste besteht darin, daß in einer Sprache die physische Begriffsreihenfolge mit dem logischen Aufbau ihrer Sätze übereinstimmt. Daraus folgt, daß logisch Zusammenhängendes auch sprachlich eindeutig als zueinandergehörig angesehen werden muß. Die zweite Feststellung besteht in der Erkenntnis, daß logische Beziehungen in einer Sprache durch horizontale und vertikale Strukturen zum Ausdruck gebracht werden können.

Ohne bislang auf die Belange der elektronischen Datenverarbeitung eingegangen zu sein, erhebt sich hier die Frage, inwieweit diese sprachanalytisch abgeleiteten Ergebnisse im Rahmen der Datenorganisation verwirklicht sind oder ob, warum und inwieweit von diesen Ergebnissen abgewichen werden muß.

2.1.2 Logische Segmentierung

Ziel der Datenorganisation ist die Schaffung von Zugriffsmöglichkeiten zu Daten, die möglichst weitgehend unabhängig von der physischen Speicherung dieser Daten auf Datenträgern sein sollen. Dieses gilt insbesondere für den Aufbau einer Datenbank, da hier das Kriterium der Benutzerorientiertheit im Vordergrund zu stehen hat.

Bei Datenbanken wird die Datenkonzeption des Benutzers als Datenstruktur (formale Datenorganisation), die physische Speicherung eines Datums auf einem Datenträger als Speicherstruktur (physische Datenorganisation) bezeichnet.¹ Im Gegensatz zur speicherorientierten physischen Datenorganisation werden die Datenstrukturen, die bei der formalen oder logischen Datenorganisation von Interesse sind, primär durch benutzerorientierte Datenbeschaffungs- und Fortschreibungsprozesse geschaffen.²

Die Unterscheidung in physische und logische Sätze ist bei der Datenorganisation von besonderer Bedeutung. Der formale oder logische Satz besteht aus Segmenten, die allein aufgrund logischer Gesichtspunkte gebildet werden. Segmente sind die Zusammenfassung von Feldern, die eine zusammengehörige logische Einheit bilden und die unter ähnlichen Ordnungskriterien auffindbar sind. Die Felder eines Segments, die mit einem oder mehreren Namen versehen sein können, werden entsprechend der darin enthaltenen Datenelemente in unterschiedlicher Häufigkeit zur Verarbeitung herangezogen. Diese Verarbeitungshäufigkeit, über die Statistiken geführt werden, entscheidet darüber, welche relativ selten benutzten Felder eliminiert werden sollen,

1 Vgl. CODASYL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, New York 1971, S. 463 ff.

2 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 78ff.

um Speicherkapazität und Speicherkosten einzusparen.¹

Ein Datenfeld wird durch vier Merkmale charakterisiert, und zwar durch den Feldnamen, die Feld-Wertmenge, seine Funktionen und sein Format. Jedes Feld muß über einen Namen verfügen, um es abfragen zu können. Unter der Bedingung, daß ein Satz in einer gegebenen Abfolge von Speicherworten abgespeichert wird, ist der Name des Feldes die Speicheradresse. Die Feldwerte sind die in einem Feld enthaltenen codierten Verweise zu einem anderen Satz oder zu einem anderen Segment. Schließlich bestimmt die Größe eines Feldes sein Format.²

Diese formal orientierten Segmentierungen lassen sich durch Bildung komplexerer Einheiten - wie z.B. eine Datei - durchaus erweitern. Sätze, die die gleiche Struktur oder den gleichen Beziehungszusammenhang aufweisen, lassen sich zu Dateien zusammenfassen. Das wesensbestimmte Element der Datenverarbeitung liegt darin, daß der benutzerorientierte logische Satz aus nach logischen Gesichtspunkten zusammengestellten Segmenten besteht, deren Abfolge nicht mehr mit der physischen Reihenfolge auf dem Speichermedium übereinzustimmen braucht. Die Zusammenfassung von Segmenten innerhalb eines logischen Satzes kann hierarchischer Natur sein, während dieses für die physische Reihenfolge nicht mehr zu gelten braucht. Insbesondere bei höherwertigen Datenstrukturen kann es aber durchaus sein, daß ein logischer Satz aus Segmenten mehrerer physischer Sätze besteht³, und die physischen Sätze sich

1 Benner hat eine Methode zur Ermittlung und Eliminierung relativ überflüssiger Felder entwickelt. Vgl. hierzu Benner, F.H., On designing generalized file records for management information Systems, in: AFIPS Conference Proceedings, Vol. 31, 1967, S. 291ff.

2 Vgl. hierzu Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 74

3 Zur Unterscheidung zwischen statistisch und dynamisch logischen Sätzen vgl. Lutz, Th. / Klimesch, H., Management Information Systems (MIS) II - Die Datenbank und ihre Probleme, in: IBM Nachrichten, 18. Jg., 1968, S. 457ff, hier S. 458f.

dadurch auszeichnen, daß zwischen ihrer hierarchischen Struktur und der tatsächlichen Speicherungsabfolge auf einem Datenträger - unter Abweichung von der obigen Feststellung - Übereinstimmung besteht.

Durch das Auseinanderfallen von logischer Abfolge von Segmenten und physischer Reihenfolge auf einem Speichermedium ist das sprachanalytisch abgeleitete Ergebnis nicht immer zutreffend, daß die logische Reihenfolge von sprachlichen Sätzen mit der physischen Reihenfolge ihrer Begriffe übereinstimmt.

Um die logisch orientierten Segmentierungsbemühungen zu erleichtern, müssen an den zu segmentierenden Datenbestand bestimmte Anforderungen gestellt werden. Ziel ist es, den Datenbestand durch Bildung logischer Segmente und Sätze in eine konsistente Ordnung zu bringen, die einen einfachen und schnellen Zugriff zu den einzelnen Datenelementen ermöglichen soll. Diese Anforderungen können dadurch erfüllt werden, daß die Anzahl der in einem Datum residierenden Ordnungskriterien beachtet wird. "Daten als Grundelemente numerischer Datenbanken lassen sich nämlich als Werte definieren, ...die mit Zuordnungskriterien ...versehen sind, so daß man über diese Zuordnungskriterien die Werte innerhalb der Datenverarbeitung verarbeiten kann."¹

Dieses soll anhand eines aus dem Marketing entlehnten Beispiels verdeutlicht werden. Die Ordnungskriterien bei der Vertragsabschlußsumme eines Vertreters können z.B. seine Personalnummer, seine Auftragskennziffer und die Artikelnummer sein. Die folgende Abbildung besagt, daß der Vertreter mit der Personalnummer 5 seinen 66-ten Auftrag über den Artikel 1002 mit einem Verkaufspreis

¹ Lutz, Th. / Beutler, H., Management Information Systems (MIS) - Begriffe und Konzeptionen für Management-Informationssysteme, in: IBM Nachrichten, 18. Jg., 1968, S. 367 ff, hier S. 367.

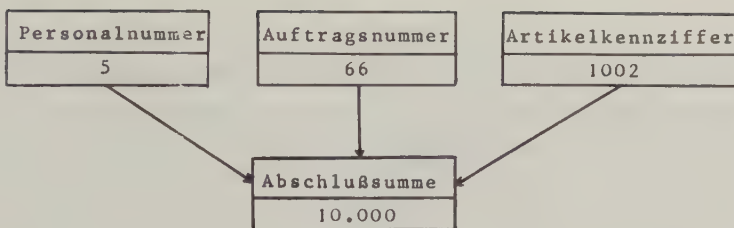


Abb. 4: Logische Ordnungskriterien

in Höhe von 10.000 DM abgeschlossen hat. Die Eignung der Datenbank für die Lieferung bedeutsamer Management-Informationen ist um so eher gewährleistet, wenn der Zugriff zu den gespeicherten Daten nach sämtlichen in ihnen residierenden Ordnungskriterien ermöglicht wird.

Darüber hinaus führt die Berücksichtigung von Ordnungskriterien dazu, daß die Kosten der Datenspeicherung gesenkt werden und die Redundanzgefahr vermieden wird, daß "Zahlenfriedhöfe" entstehen, die aufgrund fehlender Ordnungskriterien für eine effiziente Auswertung des Datenbestandes ungeeignet sind. Letztlich ist die datenverarbeitungsspezifische Festlegung eines Datums durch die Anzahl seiner immanenten Ordnungskriterien entscheidungsbestimmt und wird durch die Frage festgelegt, welche Daten überhaupt in die Datenbank eines Marketing-Informationssystems aufgenommen werden sollen. Diese Frage läßt sich generell nur unter Bezugnahme auf die Anzahl der in einem Datum residierenden Ordnungskriterien entscheiden: Mit zunehmender Anzahl der in einem Satz residierenden Ordnungskriterien nimmt der "Wert" dieses Satzes für die Datenverarbeitung zu¹, so daß die These plausibel erscheint, daß Daten mit vielen Zuordnungskriterien längere Zeit in der Datenbank gespeichert werden. Die meisten Zuordnungskriterien betreffen dabei "Stammdaten", die "Lexikoncharak-

¹ Vgl. Lutz, Th. / Beutler, H., Management Information Systems, a.a.O., S. 367.

ter" haben und mit geringfügigen oder keinen Änderungen permanent in der Datenbank verbleiben. Bewegungsdaten werden dagegen relativ häufig geändert. Die Länge der Aufbewahrungszeit von Daten ist somit Ersatzkriterium für die in den Daten residierenden Ordnungskriterien und somit für deren "Wert" im Rahmen der Datenverarbeitung.

Es sind jedoch nicht ausschließlich für Management-Informationen optimale Zugriffsmöglichkeiten zu den Datenbeständen erforderlich, sondern auch auf dem Gebiet der periodischen Massenverarbeitungen werden optimale Zugriffsmöglichkeiten zu den gespeicherten Beständen benötigt. Das Grundproblem für die Schaffung von Zugriffsmöglichkeiten läßt sich durch folgende Frage zum Ausdruck bringen: "Wie soll man die in einer Organisation auftretenden Datenelemente festlegen, gruppieren und schließlich zu Dateien zusammenfassen, so daß daraus nachher eine Datenbank entsteht, die im Hinblick auf die Manipulierbarkeit ihrer Elemente optimal ökonomisch strukturiert ist?"¹

Die gespeicherten Datenbestände müssen darüber hinaus fortgeschrieben und ggf. erweitert werden. "Kein vielgepriesenes MIS ist denkbar, wenn nicht die Grunddaten... vollständig, richtig und zweckmäßig geordnet und zeitgerecht vorliegen."² Die Existenz von Massenverarbeitungen, verbunden mit dem Bedürfnis nach Management-Information ist unabhängig von der dem Informationssystem zugrundeliegenden Struktur des sozio-technischen Systems. Aus diesem Grund gilt die Forderung nach optimalen Zugriffsmöglichkeiten für alle Datensammlungen, folglich auch für die eines Marketing-Informationssystems.

Mit der formal-logischen Segmentierung der Datenbestände entsprechend der in ihnen residierenden Ordnungskriterien

1 Lindemann, P., Die Organisation des Rechnungswesens bei automatisierter Datenverarbeitung, in: IBM Nachrichten, 18. Jg., 1968, S. 177ff, hier S. 180.

2 Lutz, Th., Was ist eine Datenbank?, in: Bürotechnik und Automation, 8. Jg., 1967, S. 250ff, hier S. 251.

sind erste Ansatzpunkte für die Erzielung optimaler Zugriffsmöglichkeiten geschaffen. Entsprechend der Fülle möglicher Ordnungskriterien kann diese logische Segmentierung äußerst unterschiedlich vorgenommen werden. Anhand eines Beispiels eines Artikelstammsatzes soll dieses näher verdeutlicht werden.

Im Artikelstammsatz wird der Verkaufspreis des Artikels, der Umsatz des laufenden Jahres und der des Vorjahres und die Nachfrage zusammengefaßt. Bei der Nachfrage wird zwischen bestellten und gelieferten Mengen unterschieden. Die in folgender Abbildung dargestellte hierarchische Struktur wird als "logischer Baum"¹ bezeichnet.

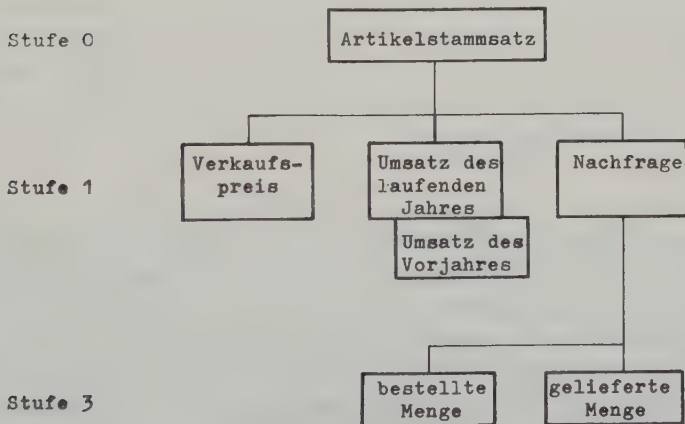


Abb. 5 : Logischer Baum

Die Kennzeichnung der hierarchischen Abstufungen erfolgt bei einem logischen Baum durch fortlaufende Nummerierung,

¹ Vgl. hierzu Wedekind, H., Datenorganisation, a.a.O., S. 39ff.

wobei das Hauptsegment die Stufe Null, das unterste Segment die höchste Stufennummer erhält. Ein logischer Baum kann als Spezialfall eines linearen Graphen¹ angesehen werden. Er ist ein endlich verbundener Graph ohne Kreise, der über mindestens zwei Knoten verfügt.² Das Hauptsegment (oder Kopf- oder Wurzelsegment) ist dadurch gekennzeichnet, daß ein "Knoten" im Sinne der Graphentheorie vorliegt, der keine hineinführenden Kanten hat, während ein Segment dann als "Blatt" bezeichnet wird, wenn es keine herausführenden Kanten hat. Sämtliche Knoten, die mit dem Wurzelsegment durch einen Pfad verbunden sind, werden als "Abkömmlinge" des "Elternknotens" bezeichnet. In der obigen Abbildung besitzt die Stufe Null ein einziges Segment - den Artikelstammsatz - während auf der ersten Stufe drei verschiedener Segmenttypen auftreten, von denen das Segment "Umsatz" zweimal auftritt. Das Segment "Nachfrage" ist isoliert für sich betrachtet Kopfsegment der beiden untergeordneten Segmente "bestellte Menge" und "gelieferte Menge". Eine derartige Struktur wird als "Unterbaum" bezeichnet.

Problem ist, inwieweit eine derartige Struktur als linear oder nicht-linear bezeichnet werden kann. Linearität ist nach WEDEKIND³ dann gegeben, wenn jedes Segment - abgesehen vom Kopfsegment und den Segmenten mit den höchsten Stufennummern - eindeutig durch vor- und nachgelagerte Segmente festgelegt werden kann. Folgt man dieser Auffassung, so ist der dargestellte logische Raum nicht-linear.

Entgegengesetzter Auffassung ist LUTZ⁴, für den die obige hierarchische Struktur linearen Charakter hat. Nichtlinea-

1 Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 39.

2 Zu den Eigenschaften von Bäumen vgl. Sussenguth, H.E., Use of Tree Structures for Processing Files, in: Communications of the Association for Computing Machinery, New York 1963, S. 272ff.

3 Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 199.

4 Vgl. Lutz, Th., Formatierte Datenbanksätze, in: Bürotechnik und Automation, 10. Jg., 1969, S. 48ff, hier S. 49ff.

rität ist nach seiner Auffassung dann gegeben, wenn Querverbindungen zwischen hierarchischen Strukturen bestehen. Der Begriff der Linearität wird im folgenden im Sinne von bestehenden Querverbindungen zwischen den Segmenten verwendet, weil erst die Errichtung dieser Querverbindungen dem Marketing-Management die Möglichkeit bietet, logische Sätze mittels der Datenbank zu generieren.¹ Aus einer Baumstruktur obiger Art folgt, daß Angaben, die für sämtliche Segmente maßgebend sind, nur einmal im jeweiligen Kopfsegment gespeichert zu werden brauchen. Eine derartige logische Segmentierung dient nicht nur der Förderung optimaler Zugriffszeiten, sondern trägt auch aufgrund der Verringerung von Datenredundanzen dazu bei, die Speicherkosten zu senken.

Mit der Gruppierung von Datenelementen in hierarchische Ordnungen ist jedoch das Zugriffsproblem keineswegs vollständig gelöst.² Vielmehr besteht eine gewisse Diskrepanz zwischen den Anforderungen der operierenden Ebene mit ihren Massenverarbeitungen und den Auswertungsanforderungen des Managements zur Erzielung von Management-Informationen.

Während auf operierender Ebene durchaus Möglichkeiten zur hierarchischen Gruppierung von Datenelementen bestehen, stehen die Auswertungen des Managements unter dem Primat der Bildung umfassender logischer Sätze. Dem Erfordernis nach primär hierarchischen Ordnungen der Datenbestände für die Belange der operierenden Ebene steht die Forderung nach Verknüpfungsmöglichkeiten der Segmente verschiedener physischer Sätze bzw. Segmente aus unterschiedlichen Dateien mit dem Ziel der Schaffung von Management-Informationen z.T. konträr gegenüber. Diese Diskrepanz ist nur dann datenverarbeitungstechnisch zu lösen, wenn neben hierarchi-

1 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 87.

2 Bei der Datenverarbeitung ist zu beachten, daß in vielen Fällen der physische Satz so angelegt ist, "daß er mit dem die Hauptverarbeitung bestimmenden Satz übereinstimmt." Lutz, Th., Formatierte Datenbanksätze, a.a.O., S. 52.

schen Strukturen auch Querverbindungen möglich sind¹, die es ermöglichen, unmittelbar auf Segmente in unterschiedlichen Dateien zuzugreifen, um logisch strukturierte Sätze aufzubauen.

Es lassen sich mittels der gängigen Programmiersprachen bereits hierarchische Datenstrukturen aufbauen², so daß sich ein Konzept der Datenorganisation, daß sich auf die Schaffung hierarchischer Strukturen beschränkt, dazu führt, daß der Bereich der klassischen Datenverarbeitung nicht verlassen wird.³ Erst die Errichtung von Querverbindungen zwischen den Segmenten einer gleichen oder einer anderen Datei schafft wirkliche "Datenbankeffekte".

Dieser Gesichtspunkt läßt sich dadurch verallgemeinern, daß die Entwicklung auf nicht-lineare Strukturen gerichtet sein muß. Die Schaffung dieser erweiterten Nicht-linearität ist primär ein Problem der Adressierungstechnik. Die Schaffung solcher Querverbindungen, um optimale Zugriffsmöglichkeiten zu allen Segmenten einer Datenbank zu ermöglichen, impliziert neben der adreßtechnischen Verkettung auch die Entwicklung von komplexen Adreßstrukturen. Dieser Problemkreis soll nun im folgenden untersucht werden.

1 Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 246.

2 Vgl. dazu die Ausführungen über die Programmiersprachen bei CODASL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, a.a.O., S. 460ff.

3 Vgl. Lutz, Th., Formatierte Datenbanksätze, a.a.O., S. 54.

2.2 Physische Datenorganisation

2.2.1 Indizierung von Datenelementen

Das Grundproblem der physischen Datenorganisation besteht darin, Datenelemente wiederaufzufinden, die - obwohl sie logisch zusammengehören - in unterschiedlichen Segmenten gespeichert sind. Dieser Problemkreis betrifft in erster Linie die Indizierung von Datenelementen.¹

Die Aufgabe der Indizierung besteht darin, daß einem Ordnungskriterium, das für ein Datenelement steht, eine Adresse zugeordnet wird, unter der ein Datenelement innerhalb einer fortlaufenden Menge von Datenelementen zu finden ist.² Unter der Adresse mittels der ein Segment auf einem Datenträger erreicht werden kann, versteht man die datenverarbeitungstechnische Entsprechung seines Ordnungskriteriums.³ "Um Datenstrukturen informationstechnologisch organisieren, speichern und auf dem laufenden halten und abrufen zu können, ist es notwendig, mit Hilfe von Indexstrukturen die Brücke zu den Speicherstrukturen zu schlagen, d.h. auf die Speicheradresse der jeweiligen Daten hinweisen zu können."⁴ Die Adresse gibt dabei den tatsächlichen Platz des Datenelements auf dem Datenträger an. Dies impliziert, daß die physische Datenorganisation konzeptionell der jeweiligen Hardware des Datenträgers angepaßt sein muß. Der Zusammenhang zwischen einem Ordnungskriterium und der zugehörigen Position des Datenelements auf dem Datenträger - die Adresse - wird recht anschaulich als Zeiger oder Index bezeichnet.⁵ Ein Index ist

1 Vgl. Dworatschek, S., Einführung in die Datenverarbeitung, Berlin - New York 1971, S. 56.

2 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 102.

3 Vgl. Lutz, Th., Adreßketten und Datentechnik, a.a.O., S. 180

4 Will, H.J., Datenbank-Systeme, in: Zeitschrift für Betriebswirtschaft, 41. Jg., 1971, S. 815ff, hier S. 825.

5 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 102.

eine Menge von codierten Informationen, die zur Wiederauf-
findung von Datensätzen aus einer oder mehreren Dateien
verwendet werden.

Es sind bei den Indexstrukturen eine Reihe verschiedener
Typen entwickelt worden, deren geeignete Auswahl sowohl
zu einer Verkürzung der Verarbeitungszeiten als auch zu
einer Verringerung des Speicherplatzes und damit der
Speicherkosten führt.

Gelingt es, von einem bestimmten Ordnungskriterium auf
verschiedene Adressen zu schließen, so liegt eine mehr-
dimensionale Indizierung¹ vor. Idealerweise sollte die
mehrdimensionale Indizierung derart beschaffen sein, daß
über ein Ordnungskriterium auf Segmente in unterschiedli-
chen Dateien zugegriffen werden kann. Können mehrere Seg-
mente verschiedener Dateien durch Zeiger verbunden werden,
ist eine Querindizierung gegeben. Schließlich liegt eine
Vielfachindizierung dann vor, wenn einer bestimmten Adres-
se mehrere Ordnungskriterien zugeordnet werden können. In
diesem Fall kann das Hauptsegment, wenn mehrere Unterseg-
mente vorhanden sind, durch diese Untersegmente erreicht
werden.

Der "Datenbankeffekt" einer Datei wird um so größer sein,
je unabhängiger von ihrer physischen Organisation die Ver-
arbeitungsmöglichkeiten dieser Datei gestaltet werden kön-
nen. Sowohl die Verarbeitungs- als auch die Organisations-
formen werden durch die Art des zur Anwendung gelangenden
Speichermediums bestimmt. Aufgrund der oft angewandten
Bandspeicherung ist die physische Organisation oft Hemmnis
bei der Verwirklichung umfassender logischer Organisatio-
nen. Eine Klassifizierung der innerhalb der physischen
Datenorganisation zur Anwendung gelangenden Speichermedien
erfolgt nach der Art des Zugriffs zu den Datenbeständen.

¹ Vgl. hierzu Lutz, Th. / Klimesch, H., Management Infor-
mation Systems, a.a.O., S. 458.

Die Literatur unterscheidet Medien mit indirektem, wahlfreiem und direktem Zugriff.¹

Das Magnetband² ist ein Speichermedium, bei dem - technisch bedingt - nur ein indirekter Zugriff möglich ist. Ein Zugriff zu einem Satz auf diesem Band ist nur dann gegeben, wenn alle physisch vor ihm auf dem Datenträger gespeicherten Sätze seriell verarbeitet worden sind. Bei Speichermedien mit wahlfreiem Zugriff wird fakultativ direkt oder indirekt zugegriffen. Für Datenbanken, die insbesondere "höherwertige" Indizierungen ermöglichen sollen, kommen nur Speichermedien mit wahlfreiem Zugriff wie die Magnetplatte oder die Magnettrommel in Frage. Aus Gründen der Speicherökonomik ist es jedoch wenig sinnvoll, den Hauptspeicher für eine permanente Datenspeicherung einzusetzen, da die beschränkte Speicherkapazität und die damit verbundenen Speicherkosten dazu in Widerspruch stehen. Eine effiziente, d.h. auch ökonomisch vertretbare Speicherung von Datenelementen, ist nur auf externen Speichern gewährleistet.

Für die physische, datenträgerorientierte Organisation formal gegliederter Datenbestände sind spezielle Verfahren entwickelt worden, die hier unter dem Aspekt der adreßtechnischen Verknüpfung erörtert werden sollen. Bei Datenbankkonzeptionen sollen starre Verarbeitungsformen, wie sie in der traditionellen Datenverarbeitung zur Anwendung gelangen, möglichst vermieden werden, um relativ unabhängig von der physischen Organisation die Verarbeitungsmöglichkeiten in den Dateien der Datenbank gestalten zu können. Aus diesem Grund werden die Verfahren, die überwiegend in der traditionellen Datenverarbeitung angewandt werden, nur gerafft behandelt.³ Die gestreut gespeicherte und die sequentielle Organisation sind jene Verarbeitungs-

1 Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 50.

2 Vgl. zu den Speichermedien Schulz, A., Strukturanalyse der maschinellen betrieblichen Informationsverarbeitung, Berlin 1970, S. 152.

3 Vgl. hierzu Wedekind, H., Datenorganisation, a.a.O., S. 52ff und S. 84ff.

formen, die fast ausschließlich bei linearen Datenbankstrukturen Anwendung finden und eher im Bereich der traditionellen Datenverarbeitung anzusiedeln sind. Die gestreut gespeicherte Verarbeitung zeichnet sich dadurch aus, daß ein Zugriff zu den Sätzen einer Datei ermöglicht wird, ohne die Reihenfolge in der Verarbeitung dieser gespeicherten Sätze beachten zu müssen.¹ "Man kann sie nur dann erzeugen, wenn man von einem Bestand mehr weiß, als daß er nur fortlaufend gespeichert ist."²

Die sequentielle Verarbeitung ist dadurch gekennzeichnet, daß nur unter der Bedingung der Abarbeitung des dem gesuchten Segments vorgelagerten Segments ein Zugriff zu diesem möglich ist. Das Zugriffskriterium besteht lediglich in der relativen Stellung der Segmentabfolge, da keine Adresse existiert.³ Diese Verarbeitungsform wird durch die Möglichkeit bestimmt, von einem identifizierten Segment zu einem damit relational verbundenen Segment zu gelangen. Es lassen sich bei der sequentiellen Organisation zwei Arten unterscheiden, und zwar die physisch-sequentielle und die logisch-sequentielle Form. Bei der physisch-sequentiellen Organisation sind aufgrund technischer Gegebenheiten ausschließlich Abarbeitungen bis zum gesuchten Segment notwendig. Die logisch-sequentielle Form oder Listenstruktur⁴ ist

1 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 66.

2 Ebenda.

3 Vgl. CODASYL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, a.a.O., S.463.

4 Die Entwicklung von Listenstrukturen basiert auf Arbeiten von Newell / Simon und Shaw, die zur empirischen Überprüfung ihrer Problemlösungstheorie auf die Entwicklung spezieller Programmiersprachen zur Lösung von Problemlösungsprogrammen angewiesen waren. Vgl. zu den Listenstrukturen insbesondere Shaw, J.C. / Newell, A. / Simon, H.A. / Ellis, T.O., A Command Structure for Complex Information Processing, in: Proceedings of the Western Joint Computer Conference, Los Angeles 1958, S. 119; ferner Newell, A. / Tonge, F.M., An Introduction to Information Processing Language-V, in: Communications of the Association for Computing Machinery, New York 1960, S.205ff.

ne Kombination zwischen geketteter Organisation und der Anwendung von Indextabellen. Ein Listenelement besteht aus einem Datenelement und einem Zeiger und/oder Anker. In dieser geketteten Organisation verweist der Zeiger oder Anker auf das nächste Listenelement. Das Datenelement enthält entweder Felder der in der Listenstruktur organisierten Segmente oder die Adresse des ersten Listenelements einer Teilliste. Die logisch-sequentielle Form ist allein durch die logisch zulässige serielle Abfolge der Sätze gekennzeichnet. Listenstrukturen zeichnen sich besonders dadurch aus, daß spezielle Programmiersprachen zur Anwendung gelangen, um sich relativ stark ändernde Datenstrukturen wirtschaftlich organisieren zu können.¹ Die Besonderheit dieser Programmiersprachen besteht darin, daß die verwendete Listenstruktur unabhängig von der physischen Speicherung ist.²

Bei jeder sequentiellen Organisation muß das System in der Lage sein, anzugeben, wann das Ende der Sequenz erreicht ist. Ein Weg, um eine logisch-sequentielle Organisation aufzubauen, besteht in der Erstellung einer physisch-sequentiellen Liste der unmittelbar verfügbaren Adressen.³

Von den Verfahren, die für Datenbanken geeigneter sind, sind vor allem diejenigen Indizierungsverfahren zu nennen, bei denen auch die Nichtlinearität von Datenbankstrukturen hinsichtlich optimaler Zugriffsmöglichkeiten adreßtechnisch berücksichtigt werden kann. Insbesondere sind hier die indexsequentielle und die verkettete Organisationsform mit Unterformen zu nennen.

1 Vgl. hierzu Borrow, D.G. / Raphael, B., A Comparison of List-Processing Languages, in: Communications of the Association for Computing Machinery, New York 1964, S. 231ff; Henry, W.R., Hierarchical Structure for Data Management, in: IBM Systems Journal, Vol. 8, No. 1, 1969, S. 2ff; IBM, Introduction to List Processing Facilities of PL/I, IBM Formular Nr. 20-0015.

2 Vgl. zu den Programmiersprachen Shaw, J.C. / Newell A. / Simon, H.A. / Ellis, T.O., A Command Structure for Complex Information Processing, a.a.O., S. 120.

3 Vgl. CODASYL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, a.a.O., S. 465.

Bei der indexsequentiellen Organisationsform¹ werden die Datenbestände sequentiell nach einem Ordnungskriterium auf dem Datenträger geordnet. Um die Position eines Datenelements lokalisieren zu können, stehen Indextabellen zur Verfügung, die ein schnelles Auffinden der gesuchten Segmente ermöglichen. Unter einer Indextabelle wird ein Verzeichnis der Ordnungskriterien einer Datei mit den ihr zugehörigen Speicheradressen verstanden.² Diese Organisationsform ist nur bei Speichergeräten mit wahlfreiem Zugriff anwendbar. Die Indextabelle ist eine von der Datei physisch "separate Datei", die dafür zu sorgen hat, daß die Segmente wirtschaftlich wiederaufgefunden werden können. Wird das einer Indextabelle zugrundeliegende Ordnungskriterium für die Festlegung der Speicherabfolge der Sätze verwendet, so spricht man von einer Daten-Primärindextafel.³ Unter der Bedingung, daß jedem Segment in der Datei nur ein Index zugeordnet wird, liegt ein eindeutiger Index vor. Der in der Daten-Primärindextafel enthaltene Primärindex ist derjenige, der am häufigsten zur Identifizierung des verwendeten Ordnungskriteriums verwendet wird.⁴ Dieser Index sollte in den meisten Fällen eindeutig sein. Es kann aber sein, daß dem Primärindex Sekundärindizes hinzugefügt werden müssen, wenn eine eindeutige Identifizierung des Segments bei nicht eindeutigem Primärindex gewährleistet sein soll. Darüber hinaus können künstliche Primärschlüssel eingeführt werden. Eine derartige Vorgehensweise erlaubt dann eine direkte Adressierung. Sie erweist sich aber oft bei der Anpassung der Datenbank (Datenfortschreibung) als inflexibel, so daß ihr

1 In der angloamerikanischen Literatur wird diese Organisationsform als Indexed Sequential Access Method (ISAM) bezeichnet. Vgl. dazu IBM, Datenorganisation auf Speichereinheiten mit direktem Zugriff, IBM-Formular Nr. 79966-1, o.O. | Sindelfingen | 1968, S. 65ff.

2 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 103.

3 Vgl. Lutz, Th. / Klimesch, H., Management Information Systems, a.a.O., S. 460.

4 Vgl. Buchholz, W., File Organization and Addressing, in: IBM Systems Journal, Vol. 2, 1963, S. 86ff, hier S. 87.

"Wert" für nicht-lineare und fortzuschreibende Dateien in Frage gestellt werden muß. Zusammenfassend betrachtet, verfügt die indexsequentielle Organisationsform über folgende Vorteile:

1. Datenbank Anpassungen berühren bei neu aufzunehmenden Daten ausschließlich die Indextabelle und führen somit zu keinen Veränderungen innerhalb der Dateien.
2. Die Suche nach einem Segment wird gegenüber sequentiellen Organisationsformen abgekürzt. Dieses gilt um so mehr, je stärker die Indextabelle in Relation zu den Ordnungskriterien der Segmente steht. Der relationale Bezug der Indextabelle findet seine Grenze aber in der Anzahl möglicher empirischer Relationen zwischen den Segmenten.
3. Durch die Berücksichtigung sekundärer Indextabellen - und wenn möglich, ist für jedes in den Segmenten residierende Ordnungskriterium eine Indextabelle anzulegen - kann jeder sekundäre Index auf die Speicheradresse des gesuchten Segments verweisen. Diese Verweisung kann auch auf den Primärindex erfolgen, der dann auf das gesuchte Segment verweist. Liegt eine derartige Indextabelle (Indextafel) vor, ist für eine Datei die Vielfachindizierung verwirklicht.

Ein wesentlicher Fortschritt in der physischen Datenorganisation sind verkettete Organisationsformen¹. Verkettete Organisationsformen sind durch die Existenz von Zeigern charakterisiert, d.h. von Feldern, die jedem Segment zugeordnet sind und die die Adresse des logisch folgenden Segments enthalten. Die Vergabe dieser Adressen erfolgt aufgrund eines in den Segmenten residierenden Ordnungskriteriums, das über die physische Abfolge der Segmente Auskunft gibt. Bei der Verkettung ist das eingangs abgeleitete Untersuchungsergebnis erfüllt, daß für die Rea-

¹ Vgl. hierzu Wedekind, H., Datenorganisation, a.a.O., S. 72ff.

lisierung effektiver Datenbankeffekte die physische Abfolge der Segmente auf dem Datenträger nicht mehr mit der logischen Abfolge identisch zu sein braucht. Die Verarbeitung der Segmente erfolgt hier entsprechend ihrer logischen Folge, d.h. die Adresse in dem gerade verarbeiteten Segment verweist auf das nächste logische zu verarbeitende Segment. Dieser Vorgang ist in der folgenden Abbildung verdeutlicht.¹

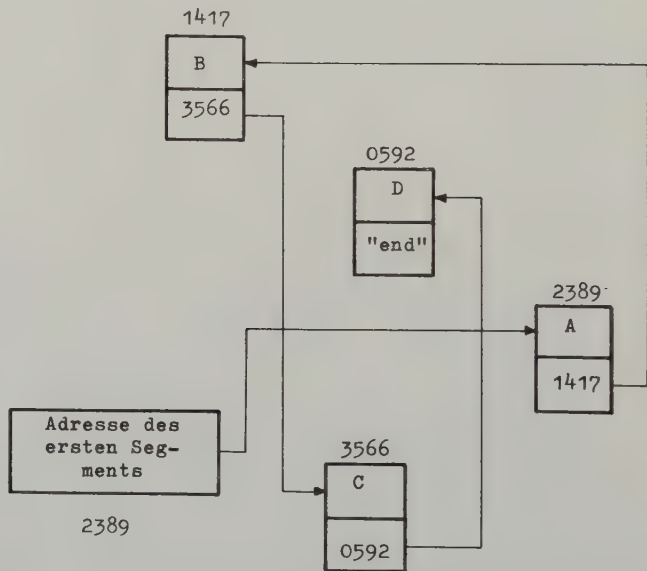


Abb. 6 : Verkettete Organisationsform

Bei der erörterten geketteten Organisation lassen sich mehrere Unterformen unterscheiden, und zwar die einfache, die doppelte und die mehrfache Verkettung. Aufgrund der

¹ Quelle: CODASYL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, a.a.O., S. 465.

logisch hierarchischen Strukturierung müssen hier entgegengesetzt der indexsequentiellen Organisation zwei Typen von Indizes unterschieden werden. Jedes Segment in der obigen Abbildung enthält nur ein einziges Adreßfeld. Dieses kann entsprechend seiner hierarchischen Stellung ein Zeiger oder Anker sein. Der Zeiger verweist immer auf ein Segment gleicher hierarchischer Stufe und Type, während ein Ankerfeld eine Adreßangabe enthält, die sich auf ein Segment unterschiedlichen hierarchischen Niveaus und/oder Typs bezieht.

Eine für Datenbanken geeignete Organisationsform innerhalb der geketteten Organisationen stellt die zyklisch gekettete Organisation dar. Bei einer zyklischen Verkettung verweist das Adreßfeld des ersten logischen Segments auf das logisch erste Segment einer Datei. Durch dieses Adreßfeld wird die Kette geschlossen, so daß es aufgrund der zirkularen Anordnung der Segmente gleichgültig ist, bei welchem Segment mit der Verarbeitung begonnen wird. Sämtliche Segmente in der Kette können durch zirkuläre Verweisungen in die Verarbeitung einbezogen werden.

Der Grundgedanke der geketteten Organisation, nämlich die Technik der einfachen Querindizierung, läßt sich durchaus erweitern. Unter Querindizierung wird die Tatsache verstanden, daß das indizierte Datenelement in einem anderen Bestand liegt als der Index. Querindizierungen können über mehrere Ordnungskriterien ausgedehnt werden, wobei diese gleichberechtigt nebeneinander als auch einander untergeordnet sein können.

Je nach Art der in den Datenelementen residierenden Ordnungskriterien lassen sich mehrfach gekettete Datei-Organisationen einführen, wobei den doppelt und dreifach geketteten Organisationen die größte Bedeutung zukommt. Der "Verkettungsgrad" wird dabei allgemein durch die Anzahl der in einem Segment enthaltenen Adreßfelder bestimmt, wobei diese Adreßfelder je nach ihrer hierarchischen Stellung sowohl Zeiger als auch Anker sein können.

Eine über diese Verkettungsformen hinausgehende Verkettung ist prinzipiell möglich, findet ihre Grenze aber in der Art der den Datenelementen gemeinsam zuzuschreibenden Ordnungskriterien oder in der Anzahl der empirischen Relationen, die zwischen den Datenelementen vorliegen können.

Die doppelt gekettete Organisation¹ zeichnet sich dadurch aus, daß jedes Segment zwei Zeigerfelder enthält, wobei das erste Adreßfeld - analog der einfachen Verkettung - auf das logisch nächste Segment verweist, während das zweite Adreßfeld die Adresse des logisch vorgelagerten Segments enthält.

Auch bei mehrwertig verketteten Organisationen sind zyklische Verkettungen möglich. Bei einer zyklischen Verkettung einer doppelt indexgeketteten Organisation verweist das erste Adreßfeld des ersten logischen Segments auf das logisch erste Segment der Datenbank, während das zweite Adreßfeld auf das logisch vorgelagerte Segment in der Datenbank verweist. Analog der zyklisch geketteten Organisation bei einfacher Verkettung ist es unbedeutend, bei welchem Segment mit der Verarbeitung begonnen wird. Darüber hinaus ermöglicht das zweite Adreßfeld bei einem vorgelagerten logischen Segment einzusetzen, wobei wiederum durch die zyklische Anordnung der Verweisungen sämtliche Segmente in die Verarbeitung einbezogen werden können. Derart gekoppelte indexgekettete Organisationen ermöglichen es, Verkettungen von Segmenten in mehreren Dateien vorzunehmen. Obwohl diese Organisationsformen für lineare Datenbankstrukturen entwickelt worden sind, lassen sich damit bereits nicht-lineare Strukturen adreßtechnisch verketteten.² Schließlich sind gekettete Organisationsformen unter besonderer Beachtung zyklischer Erweiterungsmöglichkeiten in der Lage, hohe "Datenbankeffekte" auszulösen.

1 Vgl. hierzu Wedekind, H., Datenorganisation, a.a.O., S. 82ff.

2 Vgl. Will, H.J., Datenbank-Systeme, a.a.O., S. 828.

Das Kriterium der Linearität steht aber letztlich der Realisierung dieser "Datenbankeffekte" gegenüber. Für die zu konzipierende Datenbank werden lineare Strukturen nur wenig repräsentativ sein. Vielmehr sind nicht-lineare Strukturen, wie sie beispielsweise in den Baumstrukturen der Segmente von Datenbanken ihren Ausdruck finden, typisch. Hier stellt sich insbesondere das Problem, Baumstrukturen adreßtechnisch zu verketteten, daß optimale Zugriffsmöglichkeiten zum Datenbestand gewährleistet sind. Dazu bietet sich die Anwendung von geketteten Organisationsformen in der bereits gezeigten Weise an.

Eine Organisation von Baumstrukturen kann in einfacher, zweifacher und dreifacher Verkettung vorgenommen werden.¹ Die dreifache Verkettung bietet dabei die günstigsten Möglichkeiten hinsichtlich der Zugriffszeiten und Zugriffswege. Das Grundproblem einer Baumstruktur-Organisation besteht darin, zunächst den logischen Baum derart zu strukturieren, damit der Speicherplatzbedarf auf ein Minimum beschränkt wird. Der Knoten mit dem niedrigsten Grad² folgt der Wurzel, die übrigen Knoten werden entsprechend der anwachsenden Basisgröße³ angeordnet. Ein logischer Baum ist um so eher für die adreßtechnische Verkettung geeignet, je geringer der Grad seiner Knoten in Richtung auf die Wurzel wird. In der folgenden Abbildung sind zwei logische Bäume dargestellt, von denen der erste (der ursprüngliche Baum) durch einen Umstrukturierungsprozeß in einen zweiten logischen Baum (den modifizierten Baum) überführt wurde. Bei dem in dieser Abbildung dargestellten Problem handelt es sich um ein Zugriffsstufenproblem. Es geht darum, die Zugriffsstufen zu den jeweiligen Segmenten zu verringern. Darüber hinaus

1 Vgl. zu den Verkettungsformen Wedekind, H., Datenorganisation, a.a.O., S. 200ff.

2 Unter dem Grad eines Knotens wird die Anzahl der Knoten seiner Abkömmlinge bezeichnet. Vgl. dazu den Punkt 2.1.2 dieser Arbeit.

3 Unter Basis wird die Anzahl der Abkömmlinge eines Elternknotens bezeichnet. Vgl. dazu den Punkt 2.1.2 dieser Arbeit.

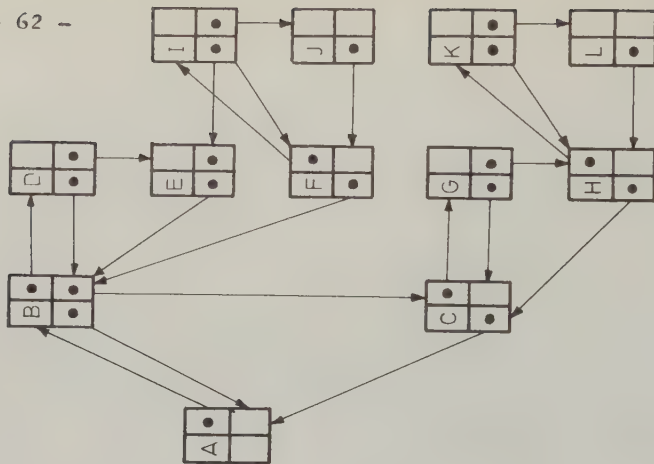
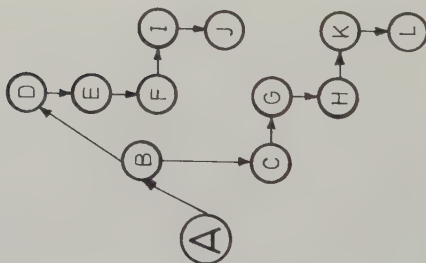
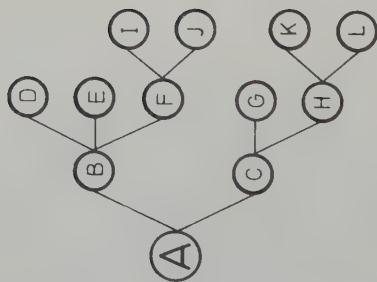
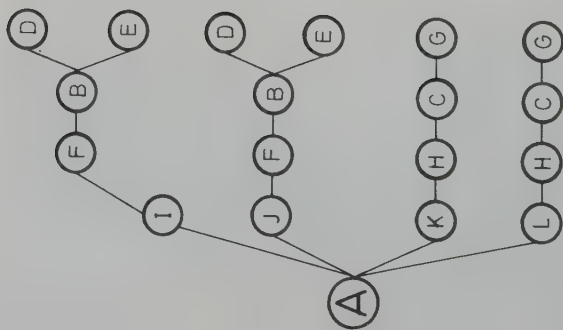


Abb. 7: Zugriffsstufenproblem

sind der Zugriffsbaum und die Speicherorganisation¹ des modifizierten logischen Baumes dargestellt, wobei die Buchstaben die jeweiligen Segmente und die Pfeile die Zugriffsrichtungen angeben.

Durch den Umstrukturierungsprozeß des ursprünglichen logischen Baumes, dessen Stufen² mit anwachsender Basisgröße zunehmen und der über 19 Knoten verfügt, ergibt sich ein modifizierter Baum mit nur 12 Knoten. Durch diese Segmentierung läßt sich der erforderliche Speicherplatz erheblich reduzieren.

Die dreifache Verkettung der Baumstruktur hat die Konsequenz, daß drei Adreßfelder benötigt werden, die entsprechend der hierarchischen Baumstruktur entweder Zeiger oder Anker sein können. Durch eine derartige Strukturierung ist es möglich, von einem zu einem anderen Segment zu gelangen, wobei der zurückzulegende Weg minimal ist. Werden derartige Verkettungen von Baumstrukturen mit Quersindizierungen zu Segmenten anderer Strukturen kombiniert, sind günstige Zugriffsmöglichkeiten gegeben.

Bei sämtlichen adreßtechnischen Verkettungen ist zu beachten, daß eine Antinomie zwischen zusätzlichen Indizierungen, die einen verbesserten Zugriff ermöglichen und dem dafür zusätzlich benötigten Speicherplatzbedarf besteht, der wiederum Einfluß auf die Speicherkosten hat. Durch Segmentierungen läßt sich Speicherplatz einsparen, wie durch Modifizierung eines ursprünglichen Baumes gezeigt werden konnte. Diese Speicherplatzersparnis kann aber ggf. durch den für die Indextabellen benötigten Speicherplatzbedarf überkompensiert werden. Dies gilt um so mehr, je komplexer die Zugriffsmöglichkeiten zu den Datenelementen einer Datenbank gestaltet werden.

¹ Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 205.

² Die Stufen eines Baumes bezeichnen den Grad der Segmentabfolge.

2.2.2 Speicherung von Datenelementen

Die Aufgabe der Speicherung von Datenelementen besteht darin, diese auf einem Datenträger zu platzieren, um sie für den Datenbankbenutzer verfügbar zu halten.¹ SCHWEIKER² verbindet mit der Speicherung ein bewußtes "Konservieren" von Datenelementen - eine Auffassung, der hier nicht zugestimmt werden kann, da der Sinn der Speicherung auch darin liegt, die Zeit zwischen einer Anfrage an die Datenbank und der Datenbereitstellung zu minimieren. Der Speicherungsprozeß als Übergang von "abstrakten" Datenelementen in ihre "materialisierte" Form ist notwendige Voraussetzung für ihre physische Manipulation.

Untersuchungen von KRAMER³ und KOSIOL⁴ auf dem Gebiet der Informations- und Kommunikationsbeziehungen zeigen, daß zwei Arten von Speicherungsprozessen unterschieden werden müssen, und zwar die Vor- und die Nachspeicherung. Die Vorspeicherung hat eine Überbrückungsfunktion zwischen dem Zeitpunkt der Erhältlichkeit von Datenelementen und ihrer Verarbeitung zu erfüllen. Diese Überbrückung ist nur dann möglich, wenn die Datenelemente auf einem (elektronischen) Medium gespeichert werden. Die Vorspeicherung gelangt insbesondere bei denjenigen Datenelementen zur Anwendung, von denen a-priori unbekannt ist, für welche Zwecke sie einmal genutzt werden sollen.

Die Nachspeicherung von Datenelementen beruht einerseits auf der Tatsache, daß diese einmal später wieder benötigt

1 Vgl. Kramer, R., Information und Kommunikation - Betriebswirtschaftliche Bedeutung und Einordnung in die Organisation der Unternehmung, Betriebswirtschaftliche Forschungsergebnisse Bd. 23, Hrsg. E. Grochla, Berlin 1965, S. 79.

2 Vgl. Schweiker, K.F., Grundlagen einer Theorie betrieblicher Datenverarbeitung, a.a.O., S. 87.

3 Vgl. Kramer, R., Information und Kommunikation, a.a.O., S. 84 und 89.

4 Vgl. Kosiol, E., Die Unternehmung als wirtschaftliches Aktionszentrum - Einführung in die Betriebswirtschaftslehre, Reinbek 1966, S. 175.

werden können, so daß ein Block von Stammdaten bestehen muß, auf den permanent zugegriffen werden kann. Andererseits ist es möglich, daß der Empfänger eines Datenelements aus technischen und/oder organisatorischen Gründen es nicht in Empfang nehmen kann, so daß die Gefahr einer Verzögerung zwischen der Beendigung des Bearbeitungsprozesses und der Abgabe des Datenelements an den nachgelagerten Informationsempfänger besteht.¹

Bei der mittels elektronischer Medien organisierten Vor- und Nachspeicherung von Datenelementen werden zwei Problemkreise berührt. Einerseits betrifft die Speicherung die Organisation der Datenelemente. Dieser Problemkreis wurde bereits im Zusammenhang mit den adreßtechnisch realisierbaren Indexstrukturen behandelt. Andererseits sind die technischen Auslegungen der zur Anwendung gelangenden Speichermedien zu berücksichtigen. Obwohl technische Gesichtspunkte der zum Einsatz gelangenden Medien nicht Gegenstand dieser Arbeit sein können, dürfen sie aufgrund des bedeutenden Einflusses, den die Technik auf die Speicherung und Wiederauffindung von Datenelementen hat, nicht völlig ausgeklammert werden.²

Nach der Auffassung von MEADOW³ stehen bei der physischen Speicherung zwei Hauptaspekte im Vordergrund, und zwar erstens das Gleichgewicht zwischen Lese- und Schreibgeschwindigkeit der Hardware und zweitens der softwaremäßig bedingte Zugang zu den gespeicherten Datenelementen. Es ist müßig, über den ersten technisch bedingten Gesichtspunkt zu diskutieren, da der insbesondere auf mechanische Hilfen angewiesene Lesevorgang oft Engpaßfaktor der Datenverarbeitung ist. Der softwaremäßig bedingte Zugang zu den Datenelementen steht in einem "Spannungsverhältnis" zwi-

1 Vgl. Kramer, R., Information und Kommunikation, a.a.O., S. 89.

2 Vgl. Meadow, Ch. T., The Analysis of Information Systems - A Programmer's Introduction to Information Retrieval, New York - London - Sydney 1967, S. 175.

3 Vgl. ebenda, S. 176.

schen technischen und physisch-logischen Datenorganisationsgesichtspunkten. Die "Güte" eines gespeicherten Datenelements wird insbesondere durch die Zugriffszeiten und Zugriffswege bestimmt. Aufgrund des o.a. "Spannungsverhältnisses" werden Zugriffszeiten und Zugriffswege unter dem Oberbegriff des technisch-physischen Datenorganisationsaspektes¹ behandelt. Das Wiederauffinden eines Datenelements darf keineswegs losgelöst von den technischen Möglichkeiten der Hardware und der physischen Datenorganisation betrachtet werden, was besonders unter dem Gesichtspunkt einer optimalen Gestaltung der Informationswiederauffindung von eminenter Wichtigkeit ist.

¹ So etwa sinngemäß Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 106.

2.3 Technisch-physische Datenorganisation

2.3.1 Zugriffszeiten

Ein wesentlicher Gesichtspunkt für die Güte einer Management-Information ist die Schnelligkeit, mit der sie den Entscheidungsträgern zur Verfügung gestellt werden kann. Trotz der hohen Verarbeitungsgeschwindigkeit der elektronischen Datenverarbeitung ist bei Anfragen an die Datenbank darauf zu achten, daß ein gesuchtes Datenelement in angemessener Zugriffszeit auffindbar ist.

Unter der Zugriffszeit soll dabei die Zeitspanne verstanden werden, die zwischen der Anforderung von Daten auf externen Speichern und deren Bereitstellung im Hauptspeicher der Datenverarbeitungsanlage verstreicht. Die Zugriffszeit zu den Datenelementen ist datenorganisatorisch und technisch determiniert und um so günstiger, je geringer diese Zeitspanne ist.

Bei den Zugriffszeiten ist zu beachten, daß nicht nur das Marketing-Management in adäquaten Zeiten auf die in der Datenbank gespeicherten Datenelemente zugreifen will. Es müssen hier auch die Verarbeitungsbelange der operierenden Ebene - in welchem Umfang z.B. detaillierte Primärdaten in die Auswertung einbezogen werden¹ - berücksichtigt werden. Bei den Arbeiten auf operierender Ebene fallen ebenfalls beträchtliche Zugriffszeiten an, so daß sie im Gesamtkomplex der Zugriffszeit beachtet werden müssen. Um sowohl die Gesamtdurchlaufzeit für Massenverarbeitungen als auch die Anforderungen des Marketing-Managements an die Zugriffszeiten zu berücksichtigen, ist es notwendig, diese beiden Gesichtspunkte miteinander in Einklang zu bringen.

¹ Vgl. Wahl, M.P., Grundlagen eines Management-Informationssystems, a.a.O., S. 23.

Im folgenden werden ausschließlich die Zugriffszeiten bei datenbankinternem Zugriff analysiert. Es wird somit von der Untersuchung der Determinanten der Zugriffszeit abgesehen, die nicht datenorganisatorisch bedingt sind, wie z.B. die technische Auslegung des Informationssystems bei Datenfernverarbeitung und die Struktur der zur Anwendung gelangenden Verarbeitungsprogramme. Die Zugriffszeit wird hier allein aus der Warte der technisch-physischen Datenorganisation betrachtet.

Die wichtigsten Determinanten der Zugriffszeit lassen sich in Hardware- und Softwarekomponenten klassifizieren. Aus diesem Grund vermischen sich bei der Erörterung von Problemen der Zugriffszeit sowohl technische als auch datenorganisatorische Gesichtspunkte. Die Hardwarekomponente der Zugriffszeit umfaßt dabei den technischen Prozeß des Datenzugriffs und die Übertragung von Datenelementen auf Terminals, von denen sie abgefragt werden können. Im Gegensatz dazu umfaßt die Softwarekomponente der Zugriffszeit den durch die Datenorganisation festgelegten Weg, um auf ein bestimmtes Segment aus einer Menge von Segmenten, die auf externen Speichern residieren, zugreifen zu können.

Die Zugriffsfrequenz und die Art und Weise der Zugriffe hängen von der Benutzerorientiertheit des Systems ab. Die Benutzerorientiertheit der Datenbank ist aber letztlich eine Funktion der realisierten Datenbankorganisation, d.h. durch Zugriffsfrequenz und Zugriffsmodi festgelegt. Da bestimmte Datenorganisationsformen an spezifische Datenträger gebunden sind, bestehen somit bei der Zugriffszeit enge Interdependenzen zwischen Speichermedien und anwendbaren Zugriffstechniken.

Die Diskussion des Problems der Zugriffszeiten kann somit nur innerhalb von Zugriffstechniken vorgenommen werden, die geeignet sind, weitgehend unabhängig von einer starren physischen Organisation auf bestimmte logische Segmente in der Datenbank zuzugreifen. Aus diesem Grund sind rein sequentielle Zugriffsmöglichkeiten für eine Datenbank untrag-

bar, da hier erst sämtliche einem gesuchten Segment vorgelagerte Segmente gelesen werden müssen. Folglich ist die sich ergebende Zugriffszeit erheblich. Daraus ergibt sich die Konsequenz, daß Datenbankdateien, die sowohl für Arbeiten auf operierender Ebene als auch für Managementauswertungen geeignet sein sollen, nicht auf Speichermedien gespeichert werden dürfen, die ausschließlich sequentielle Zugriffe - wie z.B. beim Magnetband - ermöglichen.¹

Unter dem Gesichtspunkt der Zugriffszeit ist vor allem die Magnetplatte als Speichermodell geeignet, die es ermöglicht, komplexere physische Organisationsformen zu realisieren.

Eine Möglichkeit, die Zugriffszeiten positiv zu beeinflussen, besteht in einer angemessenen Blockung der Datenbestände. Unter Blockung wird das Zusammenfügen eines/ mehrerer logischer Sätze zu einem physischen Satz verstanden.² Die Anzahl der in einem physischen Satz enthaltenen logischen Sätze wird durch den Blockungsfaktor zum Ausdruck gebracht. Die Einführung der Blockung ist mit dem Vorteil verbunden, daß eine sehr viel effizientere Nutzung des Speichermediums ermöglicht wird, da die Klüfte zwischen den logischen Sätzen entfallen, so daß sich die Zugriffszeit zu einem gesuchten Satz verringert. Bei einer Plattenorganisation sind die Sätze des Datenbestandes in der Regel in bezug auf eine Spur geblockt, um zu günstigen Ein- und Ausgabezeiten zu gelangen.³ Der Blockungsfaktor muß aber gewährleisten, daß der Inhalt einer Spur oder der Inhalt eines Blocks in den Hauptspeicher paßt.

1 Vgl. Reusch, G., Die Struktur des Programmsystems und der Datenbank für ein Informationssystem, in: ADL-Nachrichten, 13. Jg., 1968, S. 24ff, hier S. 25.

2 Vgl. Wedekind, H., Datenorganisation, a.a.O., S. 37.

3 Vgl. ebenda, S. 37f.

Diese einzuhaltende Bedingung stellt die Grenze einer effektiven Blockung dar.

Die Zugriffszeit ist dann minimal, wenn ein gesuchtes Segment mit einem einzigen Zugriff gefunden werden kann. Diese Forderung steht aber in Widerspruch zu den Zugriffsmöglichkeiten zu den in Dateien gespeicherten Segmenten. Der Zugriff zu den in den Dateien residierenden Segmenten ist dann um so besser möglich, je feiner und vielschichtiger die Daten segmentiert sind. Mit dieser feineren und vielschichtigeren Segmentierung nehmen aber auch die Zugriffswege zu den gesuchten Segmenten zu, so daß die Zugriffszeit zwangsläufig ansteigt. Es stellt sich somit für den Organisator einer Datenbank die Aufgabe, eine Abwägung zwischen der Zugriffszeit und den Zugriffswegen vorzunehmen.

2.3.2 Zugriffswege

Je zweckmäßiger die Datenelemente in einer Datenbank organisiert sind, um so geringer werden die Zugriffswege sein, die notwendig sind, um auf die Datenelemente zugreifen zu können. Unter einem Zugriffsweg soll die datenverarbeitungstechnische "Entfernung" zwischen Index und Datenbestand verstanden werden. Zugriffszeit und Zugriffsweg stehen dabei in einem sehr engen Zusammenhang. Um das Wahlproblem zwischen Zugriffszeiten und Zugriffswegen lösen zu können, kann der Gedanke Berücksichtigung finden, daß bei bestimmten Datenelementen tendenziell das Postulat nach angemessenen Zugriffszeiten im Vordergrund steht, während bei anderen Datenelementen die Forderung nach günstigen Zugriffswegen dominiert. Bei den Massendaten der operierenden Ebene wird der Gesichtspunkt der Zugriffszeit dominieren, während bei den Dateien, die spezifische Daten für die Zwecke des Marketing-Management enthalten, der Gesichtspunkt günstiger Zugriffswege im Vordergrund steht. Diese Gesichtspunkte werden von WAHL¹ und LUTZ/BEUTLER² vertreten, die vorschlagen, eine vertikale Aufteilung der Datenbank vorzunehmen und zwischen Dateien der operierenden Ebene und reinen Management-Dateien zu trennen. Dieses hat die Konsequenz, daß die bisher vertretene Auffassung, daß die Dateien sowohl den Belangen der operativen als auch denen der Managementebene genügen sollen, hier aufzugeben ist. Den Grundstock der Datenbank bilden die Dateien, die Datenelemente enthalten, die für Arbeiten der operierenden Ebene benötigt werden. Diese Datenelemente werden mittels Bridge-Programmen³ zu Management-Dateien verdichtet und in den Dateien für das Marketing-Management gespeichert.

¹ Vgl. Wahl, M.P., Grundlagen eines Management-Informationssystems, a.a.O., S. 80ff.

² Vgl. Lutz, Th. / Beutler, H., Management Information Systems, a.a.O., S. 370f.

³ Vgl. dazu Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 142 und S. 188.

Bridge-Programme verknüpfen dabei die Datenbestände und bieten die Möglichkeit zur weiteren Auswertung des Datenbestandes. "Es müssen also Operationen sein, die man nicht nur auf Zahlen und Zahlengebilde anwendet, sondern eben auf Bestände".¹ Derartige Programme arbeiten nur unter der Bedingung, daß Datenbeschreibungstabellen vorhanden sind, so daß Datenmanipulationen möglich sind. Die Konzeption der Bridge-Programme kann dahingehend erweitert werden, daß jeder Management-Ebene eine "eigene Datenbank" zur Verfügung gestellt wird.² Das Grundproblem derart "selbständiger" Datenbanken besteht darin, daß bei den verdichteten Datenelementen der Zugriff zu den Grunddaten, aus denen diese Datenbanken durch Bridge-Programme geschaffen wurden, gefährdet sein kann. Es ist also darauf zu achten, daß von allen Ebenen der Datenbank der Zugriff zu den Grunddaten möglich ist. Anderenfalls sind zeitaufwendige Sortierprozesse notwendig, die sich ungünstig auf die Zugriffswege auswirken können.

Diese vertikale Datenbankstrukturierung hat erhebliche Konsequenzen in bezug auf Zugriffswege und Zugriffszeiten. Die Zugriffszeit wird sich durch Einschalten von Bridge-Programmen zunächst erhöhen, während sich dann anschließend aufgrund verkürzter Zugriffswege ein positiver Einfluß auf die Zugriffszeit ergibt. Inwieweit sich ein positiver Nettoeffekt ergibt, kann durch einen Vergleich zwischen dem zusätzlichen Zeitaufwand für die Bridge-Programme und der aufgrund verkürzter Zugriffswege sich ergebenden Zeitersparnis ermittelt werden.

Unter der Bedingung einer gegebenen formalen und physischen Datenorganisation kann eine positive Beeinflussung der Zugriffszeit nur durch geeignete physische Datenorganisations-Verfahren erfolgen. Es kommt hier insbesondere darauf an, die physische Anordnung von Datenelemen-

1 Vgl. dazu Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 142.

2 Vgl. dazu Lutz, Th. / Beutler, H., Management Information Systems, a.a.O., S. 370f.

ten auf Datenträgern hinsichtlich der Zugriffswege günstig zu gestalten. Insbesondere bei der indexsequentiellen Datenorganisation lassen sich durch geeignete Anordnung der Dateien viele Zugriffswege einsparen, so daß sich daraus ein positiver Einfluß auf die Zugriffszeiten ergibt.

Bei der indexsequentiellen Organisation wird zunächst auf den Index zugegriffen, dieser wird sukzessive durchsucht, bis Identität zwischen dem Schlüssel- und dem Suchargument gegeben ist.¹ Über die Adresse aus dem Index wird dann auf den physischen Speicherplatz des gesuchten Segments zugegriffen. In der Regel wird der Index separat vom Datenbestand gespeichert. Eine bedeutsame Verringerung des durchschnittlichen Zugriffsweges ergibt sich dadurch, indem der Index in die Mitte des Datenbestandes verlegt wird, so daß sich der durchschnittliche Zugriffsweg gegenüber der getrennten Speicherung von Index und Datenbestand um die Hälfte reduziert. Die günstigste Möglichkeit hinsichtlich der Gestaltung der Zugriffswege bietet das Einlesen des Index in den Hauptspeicher der Datenverarbeitungsanlage vor Beginn der Verarbeitung. Dadurch entfallen, wenn dieses von Seiten der Software möglich ist, die Zugriffswege zwischen Index und Datenbestand.

Die weiterentwickelten Datenorganisationsformen - insbesondere die geketteten und hierarchisch strukturierten Datenbestände - lassen sich nur bei sehr langen Zugriffswegen verarbeiten.² Es fehlt zwar in der Literatur nicht an Vorschlägen zur Minimierung der Zugriffswege, diese Vorschläge sind aber mit der gegenwärtigen Hardware nicht zu realisieren oder sind Idealmodelle, die bewußt von den hardwaremäßigen Möglichkeiten abstrahieren, um Algorithmen zu entwickeln, die für günstige Zugriffszeiten und Zugriffswege sorgen.³

1 Vgl. dazu Wedekind, H., Datenorganisation, a.a.O., S. 60.

2 Vgl. ebenda, S. 76.

3 Vgl. Sussenguth, H.E., Use of Tree Structures for Processing Files, a.a.O., S. 272ff. Eine ausführliche Darstellung dieser Algorithmen findet sich auch bei Wedekind, H., Datenorganisation, a.a.O., S. 180ff und 256ff.

3. Verwaltung der Datenbank

3.1 Kommunikation mit der Datenbank

3.1.1 Grundprobleme der Datenbank-Kommunikation

"The objective of the data management function is to provide access to information contained in the data base. It serves as the interface between storage hierarchy and the various information processes that use or modify the contents of the data base."¹ Die Datenbank-Verwaltungsfunktionen betreffen nach dieser Auffassung die Kommunikationsmöglichkeiten mit der Datenbank, einschließlich der mit der Kommunikationssprache zusammenhängenden Hauptfunktionen² sowie die Eröffnung, den Aufbau, die Pflege und die Modifikation von Dateien.

Um die Komplexität dieser Datenbankprozesse³ zu reduzieren, ist es opportun, zwischen Prozessen zu trennen, die zu keiner Änderung des Datenbankinhalts führen und solchen, bei denen der Inhalt der Datenbank verändert wird.

Direktanfragen an die Datenbank haben keinerlei Änderungsprozesse zur Folge.⁴ Sie werden adhoc - ohne daß ihr Zeitpunkt prognostizierbar ist - an die Datenbank gestellt. Im Gegensatz dazu führen die Zusatzfunktionen der Kommunikationssprache zur Änderung eines bestehenden Datenbankvolumens.

Ein weiterer Gesichtspunkt, der in Zusammenhang mit einer durch Zusatzfunktionen ergänzten Kommunikationssprache auftritt, besteht in der Berücksichtigung von Gegebenheiten des der Datenbank abbildungsmäßig zugrundeliegenden

1 Emery, J.C., Organizational Planning and Control Systems, a.a.O., S. 49.

2 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 146ff.

3 Vgl. Will, H.J., Datenbank-Systeme, a.a.O., S. 827ff.

4 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 148; Emery, J.C., Organizational Planning and Control Systems, a.a.O., S. 50ff.

Marketing-Systems. Eine diesbezügliche Berücksichtigung ist nur über die Zusatzfunktionen möglich, bei denen die Datenbank Anpassungsprozesse dafür zu sorgen haben, daß ein bestehendes Datenbankvolumen permanent an Änderungen des Marketing-Systems angepaßt wird. Insbesondere im Zusammenhang mit der Mensch-Computer-Kommunikation haben schließlich Sicherheitsvorkehrungen dafür zu sorgen, daß Datenschutzprinzipien eingehalten werden.

Die Mensch-Computer-Kommunikation hat die Aufgabe, den unmittelbaren Datenaustausch zwischen der Datenbank und ihren Benutzern zu ermöglichen. Um derartige Kommunikationsprozesse zu bewerkstelligen, müssen Medien für die Übertragung von Zeichen zwischen Datenverarbeitungsanlagen und Benutzer vorhanden sein. WEIZENBAUM¹ hat 1966 ein Programm unter dem Namen ELIZA veröffentlicht, das versucht, eine Kommunikation zwischen Datenverarbeitungsanlage und Benutzer mittels der natürlichen Sprache zu ermöglichen. Dieses Programm basiert auf dem maschinellen Erkennen bestimmter Schlüsselworte, wobei die in das System eingehenden Daten nach diesen Schlüsselworten durchsucht werden. Obwohl dieses Programm für nicht-formatierte Datenbanken entwickelt worden ist, läßt sich seine Grundproblematik durchaus auf formatierte Datenbanken übertragen: Es wird eine Kommunikationssprache benötigt, deren Syntax annähernd mit der einer natürlichen Sprache übereinstimmt, so daß die Benutzerorientiertheit der Datenbank gewährleistet ist.

In der traditionellen Datenverarbeitung ist eine Benutzerorientierung kaum vorhanden. Es besteht eine Diskrepanz zwischen den Datenverwendern und dem Personenkreis, der effektiv auf die in einer Datenbank gespeicherten Daten zugreift. Erst durch Weiterleitungsprozesse gelangen die Daten zu ihren Verwendern, so daß die Daten bei der Übermittlung verfälscht und/oder verloren gehen können. Der

¹ Vgl. Weizenbaum, J., ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine, in: Communications of the Association for Computing Machinery, New York 1966, S. 36ff.

traditionelle Ansatz in der Datenverarbeitung zeichnet sich weiterhin dadurch aus, daß Daten in voneinander getrennten Dateien gespeichert werden, wobei in Hinblick auf Verarbeitungszwecke identische Daten durchaus in verschiedenen Dateien residieren können.¹ Daraus folgt, daß die Entwicklungsrichtung bei den Informationssystemen darauf gerichtet sein muß, ein System von Dateien zu entwerfen, das für eine Vielfalt unterschiedlicher Verwendungszwecke geeignet und den Benutzerbedürfnissen angepaßt ist. Die verarbeitungsorientierten Dateien der traditionellen Datenverarbeitung sind stets der Gefahr ausgesetzt, daß zwar bei Dateireorganisationen "neue" Dateien geschaffen werden, die aber letztlich nur in einer Umstrukturierung eines gegebenen Datenbestandes ihre Ursache haben und damit erhebliche Redundanzen aufweisen.²

Aufgrund eines ständig steigenden Datenvolumens und unter Berücksichtigung von Kostengesichtspunkten sind also Ansätze nötig, die die skizzierten Nachteile der traditionellen Datenverarbeitung vermeiden und es gestatten, zu einer Integration der Dateien zu gelangen. Dabei erweist sich insbesondere die mit "Zusatzfunktionen" ausgestattete Kommunikationssprache als geeignet, Daten abzufragen, neue Dateien einzurichten, bestehende Dateien abzuändern und zu löschen.³ Die Entwicklung von um diese Funktion erweiterte Kommunikationssprachen steht hier im Vordergrund der Bemühungen.

Diese Entwicklungsleitlinien werden von CODASYL in einem Zielkatalog zum Ausdruck gebracht. Die bedeutsamsten Ziele dieses Kataloges sind folgende:⁴

1 Vgl. CODASYL Data Base Task Group Report, New York, April 1971, S. 5.

2 Vgl. ebenda, S. 5f.

3 Vgl. Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 147.

4 Vgl. CODASYL Data Base Task Group Report, a.a.O., S. 6.

1. Die Daten müssen sich bezüglich der Benutzererfordernisse geeignet strukturieren lassen, unbeschadet der Tatsache, daß gewisse Datenelemente für heterogene Zwecke verwendet werden können.
2. Durch eine anspruchsvolle Suchlogik muß gewährleistet sein, daß die Daten wiederaufgefunden werden können, eine Anpassung der Datenbank möglich ist und die Dateien Schutz vor unbefugtem Zugriff bieten.
3. Dem Benutzer soll die Möglichkeit gegeben werden, die physische Datenorganisation zu kontrollieren, Datenstruktur-Organisationen miteinander zu kombinieren und unabhängig von der physischen Datenorganisation mit der Datenbank zu kommunizieren und die in der Methodendatenbank enthaltenen Programme mit denen in der Datenbank vorhandenen Daten zu verarbeiten.

Diese hier angeführten Subziele ordnen sich sämtlich den Oberzielen Flexibilität und Generalität der Datenbank unter.¹ Insbesondere wird von CODASYL² die größtmögliche Flexibilität bei Manipulationen von Datenstrukturen verlangt, die so generell wie möglich sein sollen, damit eine Fülle von Anwendungsfällen dadurch abgedeckt ist. Methodologisch soll der Zielkatalog dadurch erreicht werden, indem neben der Kommunikationsfunktion der Datenbanksprache die o.a. Zusatzfunktionen berücksichtigt werden sollen: "The Data Base Task Group allows the use of appropriate techniques for each aspect, while providing the discipline for interlocking the various parts into a unified whole."³

Bei der Kommunikation mit der Datenbank lassen sich zwei Benutzergruppen unterscheiden, und zwar der Benutzer, der über keine Kenntnisse in einer Programmiersprache verfügt und der in Programmierungsfragen bewanderte Benutzer. Das Grundproblem der Mensch-Computer-Kommunikation besteht da-

1 Vgl. dazu Emery, J.C., Organizational Planning and Control Systems, a.a.O., S. 57.

2 Vgl. CODASYL Data Base Task Group Report, a.a.O., S. 7 und S. 153ff.

3 Ebenda, S. 7.

rin, ob und inwieweit der erste Benutzertyp eine Programmiersprache zu erlernen hat, um mit der Datenbank kommunizieren zu können. Datenbankkonzeptionen neuerer Prägung zielen insbesondere darauf ab, dem Benutzer einen direkten Zugriff zu ermöglichen.

Die technische Seite des Zugriffsproblems betrifft die Anzahl, Art und Kombination von Terminals mit den Einheiten für die Dateneingabe, Datenausgabe und -weiterleitung. Obwohl hier überwiegend technische Gesichtspunkte - oft in Symbiose mit Datenfernverarbeitungsproblemen - im Vordergrund stehen, sollen diese Aspekte hier vernachlässigt werden.¹ Im folgenden werden ausschließlich Probleme datenorganisatorischer Art behandelt, soweit sie Direktanfragen an Datenbanken betreffen.

Im Gegensatz zu den Ansätzen der traditionellen Datenverarbeitung sind bei Direktanfragen an eine Datenbank "besonders jene Anfragen von Interesse, die nicht starr vorprogrammiert sind, und deren Programm über ein bestimmtes Codewort aktiviert wird, sondern die so beschaffen sind, daß die Dateien, an die sich eine Anfrage richtet, ebenso wie die individuellen Sätze, die zur Auswertung herangezogen werden sollen, erst zum Zeitpunkt der Anfrage spezifiziert werden."² Dieses ist Gegenstand zweier unterschiedlicher Typen von benutzerorientierten Kommunikationssprachen, die für Direktanfragen an eine Datenbank entwickelt worden sind.³

1 Vgl. z.B. Dworatschek, S., Einführung in die Datenverarbeitung, a.a.O., S. 264ff; IBM, Datenfernverarbeitung - IBM Teleprocessing, Formular Nr. 74968-1, Juli 1966.

2 Lutz, Th., Direktanfragen an eine Datenbank, in: Büro-technik und Automation, 9. Jg., 1968, S. 96ff, hier S. 96.

3 Bei der Behandlung der Kommunikationssprachen besteht das Dilemma, daß generelle Aussagen kaum gemacht werden können, sondern nur anhand ausgewählter Beispiele das Kommunikationsproblem verdeutlicht werden kann.

3.1.2 Direktanfragen an die Datenbank

Bei den mit Kommunikationssprachen arbeitenden Datenverarbeitungssystemen lassen sich zwei Typen unterscheiden, und zwar Systeme, die über in eine Host-Language eingelagerte Sprache verfügen und Systeme mit einem eigenen Sprachteil.¹

Systeme mit einem in eine Host-Language (Programmiersprache) eingelagerten Sprachteil zeichnen sich dadurch aus, daß neben der Syntax der Programmiersprache zusätzliche Statements notwendig sind, die den Benutzer der Datenbank in die Lage versetzen, mit ihr zu kommunizieren. Eine Anfrage an die Datenbank bezieht sich dabei immer auf bestimmte Dateien. Diese werden entweder explizit durch die Anfrage angefordert oder sie fließen je nach Systemkonfiguration durch Dateiverkettungen in die Anfrage ein. Die benötigten Dateien sind dabei durch standardisierte Statements zu beschreiben, so daß der Benutzer neben der Kenntnis der betreffenden Host-Language auch die Strukturierung der anzufragenden Dateien kennen muß.

Die von CODASYL vorgeschlagene Kommunikationssprache arbeitet mit COBOL-Sprachelementen. Es lassen sich zwei Sprachtypen unterscheiden, und zwar die Datenbeschreibungssprache und die Datenmanipulationssprache. Die Datenbeschreibungssprache wird dazu benötigt, um die Formate der Daten hinsichtlich der physischen Datenorganisation verwenden zu können. Sie enthält sämtliche Angaben, die zur Strukturierung und Identifizierung der Daten notwendig sind. Unter Berücksichtigung der Tatsache, daß sehr viele Benutzer in Kommunikation mit der Datenbank treten, und eine einzelne Datenbeschreibung die Verkettung mit nicht angefragten Dateien verhindern würde, ist es sinnvoll, wenn jede Datenbank in Form von Datenbeschreibungstabellen über ein Instrument verfügt, das dem Benutzer die Beschreibung der Dateien erspart. "Wie man

¹ Vgl. dazu CODASYL Data Base Task Group Report, a.a.O., S. 7f und S. 197ff.

sieht, haben Datenbeschreibungstabellen eine sehr zentrale Bedeutung für die Datenbank und ihre Benutzer. Sie stellen ...die Verbindung her zwischen dem System und dem Benutzer und garantieren die Unabhängigkeit und die Benutzerfreundlichkeit der Datenbank."¹

Die Datenmanipulationssprache, deren sich der Benutzer bedient, hat die Aufgabe, den Datentransfer zwischen dem EDV-Programm und der Datenbank vorzunehmen. Dieses benutzerorientierte Programm wird aus einer Mischung von COBOL-Statements und Sprachelementen der Datenmanipulationssprache geschrieben. Die Datenmanipulationssprache ist dazu geeignet, Datenelemente wiederaufzufinden, bestehende Dateien neu hinzuzufügen, Dateien zu modifizieren und zu eliminieren, kurzum sämtliche Eigenschaften einer mit "Zusatzfunktionen" ausgestatteten Kommunikationssprache zu erfüllen.

Die Trennung zwischen COBOL als Host-Language und der Datenmanipulationssprache ist dabei nur konzeptioneller Art.² Vom Standpunkt des Benutzers aus betrachtet, verwendet dieser ausschließlich *e i n e* Sprache, die die kombinierten Fähigkeiten der Host-Language und der Datenmanipulationssprache hat.³

Bei dem vorliegenden CODASYL-Vorschlag handelt es sich um keine Sprache, die als selbständiger Compiler implementiert ist. Obwohl der Datenbankbenutzer an das Erlernen der Host-Language gebunden ist, besteht der grundlegende Vorteil des CODASYL-Vorschlags darin, daß die Datenbank ohne umfangreiches Programm dazu gebracht werden kann, die vom Benutzer spezifizierte Anfrage oder die genannten Zusatzfunktionen zu erfüllen. Mit diesem Vorschlag ist

1 Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 142.

2 Vgl. CODASYL Data Base Task Group Report, a.a.O., S. 17.

3 Vgl. ebenda.

aber der Nachteil verbunden, daß die notwendige Flexibilität der Datenmanipulationen eines derart konzipierten Systems beschränkt ist.¹

Systeme mit einem als selbständigen Compiler implementierten Sprachteil zeichnen sich besonders durch ihre Benutzerfreundlichkeit aus, indem der Benutzer nicht an das Erlernen der Host-Language gebunden ist. Unter besonderer Berücksichtigung der in Datenbanken enthaltenen hohen Datenvolumina sind Systeme mit einem als selbständigen Compiler implementierten Sprachteil aber hinsichtlich ihrer Verarbeitungsgeschwindigkeit vermutlich langsamer als Systeme, die über einen in eine Host-Language eingebetteten Sprachteil verfügen. Dieser Gesichtspunkt der Verarbeitungs- und Anfrage-Beantwortungs-Geschwindigkeit wird insbesondere dann bedeutsam, wenn ein hohes Datenvolumen repetitiv zu verarbeiten ist.² Dieser Bedingungsrahmen ist für die täglichen Routine-Verarbeitungen typisch. Aus diesem Grund ist CODASYL zu dem Schluß gelangt, spezielle Systeme zu entwickeln, die über einen in eine Host-Language eingebetteten Sprachteil verfügen und den höheren Programmieraufwand in Kauf zu nehmen. Diese Auffassung ist insbesondere dadurch gerechtfertigt, da Datenbanken neben der Information des Managements vor allem für Arbeiten auf operierender Ebene eingesetzt werden. Als Nebeneffekte können hier die Vorteile eines mit einer Host-Language arbeitenden Systems genutzt werden, wenn es um die Informationsbedürfnisse des Managements geht.

Wird besonderer Wert darauf gelegt, daß auch Benutzer mit der Datenbank kommunizieren können, die über keine Kenntnisse in einer Programmiersprache verfügen, empfiehlt es sich, den Sprachteil der Datenbank als selbständigen Compiler zu implementieren. Derartige Kommunikationssprachen zeichnen sich dadurch aus, daß sie syntaktisch nicht

1 Vgl. CODASYL Data Base Task Group Report, a.a.O., S.8ff.

2 Vgl. ebenda, S. 5ff.

schwierig und leicht zu erlernen sind. Sprachen, die als selbständige Compiler implementiert werden, bezeichnet man als Query-Languages (Anfragesprachen).

Bei Anwendung einer Anfragesprache werden die vom Benutzer gewünschten Dateien erst zum Anwendungszeitpunkt spezifiziert. Idealerweise sollte eine Anfragesprache so beschaffen sein, daß die Anfrage des Benutzers sich auf die gewünschten Daten beschränkt. Hierbei besteht durchaus die Möglichkeit der Normierung der Informationswünsche. Jeder Datenbank-Benutzer äußert vorab seine Informationswünsche, die unter Codes tabelliert und gespeichert werden. Um eine Anfrage zu starten, braucht der Benutzer dann nur noch den Code einzugeben. Über die gespeicherte Tabelle sollte dann das Datenbanksystem die notwendigen Dateien ermitteln und die in Frage kommenden Sätze entsprechend der Abfragespezifizierung verarbeiten. Darüber hinaus müßte die Tabelle angepaßt werden, um den sich ändernden Informationswünschen der Benutzer zu entsprechen.¹ Da die Dateianforderungen von Anfrage zu Anfrage variieren können, ist es notwendig, daß die betreffenden Dateien der Datenbank permanent verfügbar sind.

¹ Vgl. IBM, Ein Weg zu einem Führungsinformationssystem, Management-Einführungsschrift, September 1970.

3.2 Anpassung der Datenbank

Bei den bisherigen Überlegungen wurde davon abstrahiert, daß die in die Datenbank aufgenommenen Daten zeitlich variant sind. Eine Analyse, die zeigen soll, wie sich die Daten verändern, muß deshalb zeitliche und sachliche Überlegungen beinhalten. Jedes Datum - mit Ausnahme der Ordnungskriterien - verfügt über einen gewissen Zustand zu einem bestimmten Zeitpunkt (Zustandsprofil).¹ Dieses Zustandsprofil ist durch seine zeitliche Ausdehnung charakterisiert und wird in horizontaler Richtung der Zeitachse durch Ereignisse abgegrenzt. Folglich ändert ein Ereignis das Zustandsprofil eines Datums oder einer Datenmenge. Ein Ereignis hat im Gegensatz zum Zustand keine zeitliche Dimension. Daraus folgt, daß die Zustandsprofile der betreffenden Daten unter Berücksichtigung des Ereigniszeitpunkts angepaßt werden müssen. Darüber hinaus müssen neue Daten aufgenommen und veraltete eliminiert werden.

Da Daten als Objekte und Relationen identifiziert wurden², tritt bei einer Anpassung der Datenbank das Problem auf, zeitlich und sachlich bedingte Änderungen von Objekten und Relationen zu berücksichtigen. Insbesondere durch die vorzunehmenden Anpassungsprozesse kann dem Postulat des managementorientierten Ansatzes nach Lieferung relevanter und aktueller Daten für das Marketing-Management entsprochen werden.

Die Anpassung einer Datenbank umfaßt dabei zwei grundlegende Problemkreise. Der erste Problemkreis betrifft die Frage, welche Anpassungsrichtungen der Datenbank möglich sind. Diese werden durch Änderungen im Realsystem³ deter-

1 Vgl. dazu den Punkt III./2.1.1 dieser Arbeit.

2 Vgl. ebenda.

3 Unter Änderungen des Realsystems werden sowohl Änderungen der Struktur als auch der Systemgrenzen verstanden.

miniert. Der zweite Problemkreis besteht in der datenverarbeitungstechnischen Realisierung dieser Anpassungsrichtungen.

Die sich innerhalb des Marketing-Systems vollziehenden Änderungsprozesse beeinflussen sowohl das Volumen der in der Datenbank vorhandenen Daten als auch deren formale Organisation. Dabei entscheidet die formale Datenorganisation darüber, ob die Anpassung der Datenbank durch Aufnahme neuer Daten nur zu einer Volumenserhöhung der in Segmenten organisierten Bestände führt oder ob neue Segmentierungen notwendig sind. Zunächst sollen die aus Veränderungen des Realsystems resultierenden Anpassungsrichtungen der Datenbank untersucht werden.

Im Marketing werden Volumens- und Strukturänderungen einerseits durch Wachstum¹ des Marketing-Managementsystems, andererseits durch Wachstum des Distributionssystems hervorgerufen. Das Wachstum des Marketing-Managementsystems führt u.a. dazu, daß sich die Informationsbedürfnisse des Marketing-Managements und die der operierenden Ebene verschieben. Werden zur Befriedigung dieser geänderten Informationsbedürfnisse neue Daten in die Datenbank aufgenommen, die in neuen Segmenten organisiert werden, wobei die Segmentstruktur der neuen Segmente mit der der alten Segmente übereinstimmt, liegt eine volumensmäßige Erweiterung der Datenbank in horizontaler Richtung vor. Der Terminus "horizontal" soll dabei zum Ausdruck bringen, daß die formale Datenorganisation bestehender Segmente durch die Neueinführung von Segmenten nicht beeinflußt wird. Die Einführung neuer Segmente ist auf allen Stufen eines logischen Baumes unter der Bedingung möglich, daß die Struktur der Arboreszenz nicht geändert wird.

Weiterhin kann es durch Wachstum des Marketing-Managementsystems notwendig werden, Daten zu neuen Segmenten zusammenzufassen, die nicht in bestehende Strukturen der formalen Da-

¹ Der Begriff Wachstum umfaßt sowohl positives Wachstum (Expansion) als auch negatives Wachstum (Schrumpfung).

tenorganisation eingeordnet werden können. Diese Wachstumsrichtung der Datenbank soll als vertikales Wachstum bezeichnet werden. Der Terminus vertikal soll dabei zum Ausdruck bringen, daß Struktur Erweiterungen notwendig sind, die Reorganisationen in der Datenbank notwendig erscheinen lassen. Bei vertikalem Datenbankwachstum erhebt sich sogleich die Frage, ob aus den zu Segmenten zusammengefaßten neu aufzunehmenden Daten eine Erweiterung der Datenbank vorgenommen werden soll, oder ob diese Segmente unter Modifizierung hierarchischer Strukturen dort eingliedert werden können. Dieses Problem läßt sich nicht ohne Bezug auf effektiv im Marketing-Informationssystem vorhandene logische Baumstrukturen beantworten. Es gilt jedoch prinzipiell, daß aufgrund sachlicher und zeitlicher Veränderungen der Daten¹ und zum Zweck der Redundanzvermeidung der gespeicherten Segmente neue Segmente nur dann zu neuen Dateien zusammengefaßt werden sollen, wenn auf sie häufig zugegriffen wird. Eine unnötige Verlagerung der Zugriffszeiten und -wege kann auf diese Weise vermieden werden.²

Weiterhin ist das Wachstum des Distributionssystems Ursache für zu vollziehende Anpassungsprozesse in der Datenbank. Da in der Datenbank eines Marketing-Informationssystems Daten über die Umwelt³ gespeichert sind, ist besonders unter Würdigung zeitlicher Gesichtspunkte die Anpassung dieser Daten zwingend notwendig. Dies gilt insbesondere unter Beachtung des Postulats des managementorientierten Ansatzes, aus Daten Informationen zu "destillieren", die prognostisch relevant sind. Die strukturellen als auch volumensmäßigen Änderungen der gewonnenen Daten bieten auch hier den Anlaß, eine vertikale und horizontale Anpassung der Datenbank vorzunehmen. Diese Anpassungs-

1 Vgl. dazu die Überlegungen bei Barnard, C., The Functions of the Executive, Cambridge 1938, S. 127f und Stachowiak, H., Denken und Erkennen im kybernetischen Modell, a.a.O., S. 109f.

2 Vgl. dazu den Punkt III./2.3.1 dieser Arbeit.

3 Vgl. dazu Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, a.a.O., S. 168ff.

richtungen sind natürlich nur dann zu realisieren, wenn die dazu notwendigen Daten auch beschaffbar sind. Die Bedeutung der Datenbeschaffung ist dabei unter besonderer Berücksichtigung der Dynamik der strukturändernden Prozesse sehr hoch einzuschätzen.

Paradebeispiel derartiger Umstrukturierungen sind Marktsegmentierungsstrategien, die dazu führen, daß sowohl in die Kundenstammdatei neue Kundenstammsätze als auch in die Artikelstammdatei Artikelstammsätze neu aufgenommen oder eliminiert werden müssen. Je nachdem wie straff der Zusammenhang zu der bereits bestehenden Kundenstammdatei oder der vorhandenen Artikelstammdatei ist, läßt sich ein mehr oder weniger starker Bezug zwischen den vorhandenen Dateien und den neu aufzunehmenden Sätzen feststellen. Analog zu den Datenbank-Anpassungsprozessen aufgrund des Wachstums des Marketing-Managementsystems lassen sich auch bei Wachstum des Distributionssystems horizontale und vertikale Anpassungsrichtungen der Datenbank unterscheiden.

Von einer Datenbank muß in diesen Situationen verlangt werden, daß sowohl vertikale als auch horizontale Anpassungen ohne völlige Reorganisationen bestehender Datenbankorganisationen vorgenommen werden können. Grundsätzlich ist zu beachten, daß die Realisierung von Anpassungsmöglichkeiten nicht unabhängig von der physischen Organisation der Datenbank ist. Insbesondere sind Zugriffsmöglichkeiten und Zugriffszeiten mit in den Kalkül aufzunehmen. Eine volumenmäßige Datenbank Anpassung darf nicht zu überproportionalen Zugriffszeiten führen. Weiterhin ist bei neuen Segmenten, die in einer neuen Datei zusammengefaßt sind, darauf zu achten, daß zu diesen Segmenten gleich gute Zugriffsmöglichkeiten geschaffen werden wie zu bestehenden Dateiorganisationen.

Bei der Anpassung der Datenbank lassen sich zwei fundamentale Prinzipien in Analogie zu Erkenntnissen auf dem Gebiet

des offiziellen Rechnungswesens ableiten, die die Anpassung der Datenbank betreffen.¹ Das erste Prinzip besagt, daß jedes Ereignis, das die in der Datenbank existierenden Daten beeinflußt, erfaßt werden muß. Das zweite Prinzip impliziert, daß zu bestimmten Zeitpunkten eine Inventur des Datenbestandes vorgenommen werden muß, die geeignet ist, die Zustandsprofile der Daten einer kritischen Revision zu unterziehen.

Auf dem ersten Fundamentalprinzip beruht jeglicher Anpassungsprozeß, so daß die Bedeutung dieses Prinzips hier nicht wieder gewürdigt zu werden braucht. Beim zweiten Prinzip stellt sich die Frage, ob Stichtagsinventuren oder permanente Inventuren besser geeignet sind, die Aktualität der in der Datenbank gespeicherten Daten zu gewährleisten. Im Gegensatz zur Stichtagsinventur der Dateien, die zeitpunktbezogene Zustandsprofile liefert, bietet die permanente Inventur die Möglichkeit, zeitraumbezogene Veränderungen der Zustandsprofile zu berücksichtigen. Unter dem Gesichtspunkt der Aktualität der Daten weist die permanente Inventur gegenüber der Stichtagsinventur erhebliche Vorteile auf. So liefert die Stichtagsinventur keine Informationen darüber, wie sich die Zustandsprofile der Daten im Zeitablauf geändert haben. Darüber hinaus kann der Inventurzeitpunkt zwischen verschiedenen Dateien variieren, so daß sich ein Vergleich der Zustandsprofile der Daten aufgrund ihres unterschiedlichen Zeitbezugs als problematisch erweist.

Die Methode der Stichtagsinventur wird sich nur bei solchen Datenbeständen eignen, deren Zustandsprofile zu bestimmten Zeitpunkten charakteristische Ausprägungen haben.

¹ So sinngemäß Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 42. Morgenstern weist besonders auf die Fehlerquellen hin, die entstehen können, wenn wirtschaftliche Beobachtungen auf einander folgenden diskreten Zeitpunkten vorgenommen werden.

Dies gilt in erster Linie für saisonal bestimmte Daten, deren Fortschreibung so lange nicht notwendig ist, wie sich die Saisoneinflüsse nicht merklich ändern.

Bei den meisten Daten wird sich jedoch die Anwendung der permanenten Inventur anbieten, weil sich ihre Zustandsprofile in einer nicht zu antizipierenden Weise ändern. Obwohl die Datenbankfortschreibung durch administrative Routinen sichergestellt sein wird, ist es Tatfrage, inwieweit es tatsächlich gelingt, sämtliche Daten anzupassen. Die reale "Normalsituation" wird vermutlich so gelagert sein, daß trotz Anwendung der Methode der permanenten Inventur eine Diskrepanz zwischen dem Zustandsprofil des Speicherungszeitpunktes und dem des Anfragezeitpunktes besteht.

3.3 Schutz der Datenbank

In der Datenbank eines Marketing-Informationssystems sind große Teile des Marketing-Datenbestandes einer Unternehmung akkumuliert, die einen angemessenen Datenschutz erfordern.¹ Die Daten sind in einer Datenbank mannigfaltigen Schädigungsmöglichkeiten ausgesetzt. Einerseits kann auf sie in unberechtigter Weise zugegriffen werden, andererseits besteht die Gefahr, daß Daten verloren gehen, verfälscht bzw. bewußt oder unbewußt vernichtet werden können. Diese beispielhaft aufgezeigten Schädigungsmöglichkeiten legen es nahe, den Schutz der Datenbank aus zwei Perspektiven zu sehen, erstens unter dem Gesichtspunkt der Datenscheidung durch den Benutzer (unbefugtes Zugreifen zu den Dateien, Verfälschen und Vernichten von Dateien), zweitens unter dem Gesichtspunkt der Schädigung durch das System selbst.

Die vom Benutzer ausgehenden potentiellen Schädigungsmöglichkeiten können dadurch verringert werden, daß nur ein begrenzter und fachlich qualifizierter Personenkreis Zugang zu den Dateien der Datenbank hat. Diese Benutzerbeschränkungen scheinen in Widerspruch zu der Schaffung von allgemeinen Zugriffsmöglichkeiten zur Datenbank zu stehen. Es ist eine zunächst paradox erscheinende Entwicklung, dem Benutzer erst durch weitgehende Loslösung der physischen von der formalen Datenorganisation Direktzugriffsmöglichkeiten zu schaffen, um dann diese Zugriffsmöglichkeiten zum Schutz der Datenbank wieder einzuschränken. Die allerdings nur scheinbare Paradoxie zwischen Datensicherheit einerseits und Zugriffsmöglichkeiten andererseits wird dadurch aufgehoben, daß innerhalb bestimmter Software-

1 Die Diskussion über Datenschutzprobleme hat in jüngster Zeit durch die Errichtungsversuche von Datenbanken auf Landes- und Bundesebene unter dem Gesichtspunkt des allgemeinen Persönlichkeitsschutzes einen aktuellen Akzent bekommen. Vgl. z.B. zu dieser Diskussion Oswald, A., Datenschutz - ein hessisches Modell, in: IBM Nachrichten, 20. Jg. 1970, S. 379ff.

Pakete Sicherheitsvorkehrungen getroffen werden können, die den Direktzugriff zur Datenbank nicht prinzipiell erschweren. Sie beeinträchtigen auch nicht die Art und Weise des zu vollziehenden Direktzugriffs, sondern nehmen ausschließlich eine Klassifikation der Dateien vor, auf die bestimmte Benutzergruppen zugreifen dürfen. Die Kommunikation mit der Datenbank als auch die Sicherheitsvorkehrungen stellen keine widersprüchlichen Ziele dar, sondern können durch die mit den Kommunikationssprachen verbundenen Zusatzfunktionen gemeinsam erreicht werden.

Das hier zunächst als Datenschutz oder -sicherheit postulierte Ziel erweist sich unter Berücksichtigung anglo-amerikanischer Literatur als unscharf. Die anglo-amerikanische Literatur unterscheidet zwischen "Security", die die Betriebssicherheit des Systems betrifft und "Privacy", die den Schutz der Datenbank vor unbefugter Benutzung betrifft.¹ Der Terminus "Privacy" soll im folgenden mit dem Begriff der "Benutzungssicherheit" gleichgesetzt werden.

Bei der Betriebssicherheit sollen nur solche Gesichtspunkte behandelt werden, die nicht allein technisch von Interesse sind², sondern auch über einen datenorganisatorischen "Kern" verfügen. Aufgabe von Betriebssicherheitssystemen ist es, dafür zu sorgen, daß Datenbestände nicht durch Bedienungsfehler oder Beschädigungen der Hard- und Software zu einem Verlust der Datenbestände führen. Eine Datenschädigungsmöglichkeit mit datenorganisatorischem Bezug ist beim permanenten Anpassen³ der Datenbank möglich. Hierbei werden gleichzeitig identische Datensätze durch zwei gleichzeitig arbeitende Programme abgeändert. Während

1 Vgl. dazu Bates, W.S., Security of Computer Based Information Systems, in: Datamation, May 1970, S. 60ff.

2 Bei der allein technisch orientierten Betriebssicherheit sei hier nur auf die Möglichkeit des Duplex-Betriebes von EDV-Anlagen hingewiesen.

3 Vgl. Moloney, R.F., New Generation EDP Control Considerations, in: Management Services, March/April 1968, S. 17ff, hier S. 17.

das eine Programm die abgeänderten Sätze schreibt, erfolgt durch das andere Programm wiederum ein Schreiben der abgeänderten Sätze, so daß die Wirkung des ersten Programms zunichte gemacht wird. Zwar ist es möglich, die Gefahr des Datenverlustes durch permanentes Anpassen mittels Softwarevorkehrungen zu verhindern¹, es können aber niemals Bedienungsfehler ausgeschlossen werden, die zu einem Datenverlust führen. Erst durch technische Vorkehrungen kann dieser Gefahr wirkungsvoll vorgebeugt werden. Um die Betriebssicherheit bei potentiellen Datenverlusten aufrechtzuerhalten, ist es notwendig, periodisch ein Duplikat des Datenbestandes zu erstellen. Diese Duplikats-Erstellung geschieht am besten täglich.

Der zweite Aspekt beim Schutz der Datenbank betrifft den Benutzungschutz.² Es werden in den Dateien einer Datenbank Daten gespeichert, die nicht für jedermann in der Unternehmensorganisation zugänglich sein dürfen, wie z.B. Kennziffern, Marktdaten etc., deren unbefugte Benutzung die Unternehmung erheblich schädigen kann. Somit ist dafür zu sorgen, daß diese Daten vertraulich bleiben, nur dem befugten Personenkreis zur Kenntnisnahme zugehen und vor Verfälschungen und Vernichtung gesichert werden können. Aus diesen Gründen ist sowohl eine Klassifizierung der schutzwürdigen Daten als auch des zugriffsberechtigten Personenkreises und der Kosten der vorzunehmenden Schutzmaßnahmen vorzunehmen.³

Unter der Erkenntnis, daß es einen absoluten Mißbrauchsschutz nicht geben kann, ist der Umfang der zu ergreifenden Schutzmaßnahmen in einem wirtschaftlich zu vertretenden Verhältnis zu halten. Die Verhältnismäßigkeit der Datenschutz-Kosten sollte sich dabei am Quotienten aus dem

1 Vgl. Moloney, R.F., New Generation EDP Control Considerations, a.a.O., S. 17ff, hier S. 17.

2 Vgl. Bates, W.S., Security of Computer Based Information Systems, a.a.O., S. 60ff.

3 Vgl. Schulze, J.H., Datenschutz in der Datenverarbeitung, in: IBM Nachrichten, 21. Jg., 1971, S. 640ff, hier S. 641.

"Wert" der zu schützenden Dateien und den dabei entstehenden Kosten orientieren. Als Orientierungshilfe lassen sich gewisse Effektivitätsstufen unterscheiden, die mit den dabei verursachten Kosten übereinstimmen.

Die erste Stufe in der Datenschutzkategorie bietet die Methode der zugewiesenen Terminals.¹ Unter der Bedingung, daß von mehreren Terminals aus auf die Datenbank zugegriffen wird, erfolgt bei dieser Methode eine Identifizierung der Terminals, die für die betreffende Anfrage verwendet worden sind. Die Identifizierung eines bestimmten Terminals geschieht durch einen Namen- oder Kennungsgeber.² Das System ordnet aufgrund dieser Namen dem betreffenden Terminal die Dateien zu, auf die von diesem Terminal aus zugegriffen werden darf. Dieses System ist somit von der Person des Benutzers unabhängig und nur auf die betreffenden Terminals bezogen. Ob eine Person berechtigt ist oder nicht, die von einem bestimmten Terminal zugreift, kann das System nicht feststellen. Diese Datenschutz-Methode ist nur bei gering schutzwürdigen Dateien angebracht.

Eine für Datenkategorien nächst höherer Schutzwürdigkeit anwendbare Methode ist die der Kennworte.³ Bei dieser Methode verfügt z.B. der Datenbankbenutzer über Schlüsselwörter, die es ihm ermöglichen, Dateien abzufragen.

Schließlich bietet sich für eine streng vertrauliche Datenkategorie die Methode "Authority Lists"⁴ an. Dabei wird an jeden dazu autorisierten Datenbankbenutzer ein Codewort vergeben, das ihn gegenüber dem System als Berechtigten ausweist. Die Prüfung dieses Codewortes erfolgt mittels

1 Vgl. Moloney, R.F., New Generation EDP Control Considerations, a.a.O., S. 18.

2 Vgl. Schulze, J.H., Datenschutz in der Datenverarbeitung, a.a.O., S. 642.

3 Vgl. Moloney, R.F., New Generation EDP Control Considerations, a.a.O., S. 17.

4 Vgl. ebenda, S. 18.

einer Benutzertabelle, die über die Person des Berechtigten, die zugriffsberechtigten Daten und Felder und die Zugriffsart Auskunft gibt. Darüber hinaus existiert eine Feldtabelle, die jedem Feld innerhalb einer Datei eine gewisse Sicherheitskategorie zugeordnet und je eine Rangstufe für Lese- und Veränderungsprozesse der Felder.¹

Will nun der Benutzer auf ein bestimmtes Feld zugreifen, müssen die Sicherheitscodes von Feld- und Zugriffstabelle übereinstimmen. Darüber hinaus ist es notwendig, daß Sicherheits-Rangstufen der Zugriffstabelle gleich oder größer als die entsprechenden Rangstufen in der Feldtabelle sind. Durch eine derartige Prozedur ist sichergestellt, daß Personen immer nur auf Felder einer bestimmten Rangstufe zugreifen können und nicht auf Felder höherer Rangstufe, für die diese Personen nicht befugt sind.

Diese Methode ist zwar sehr sicher, führt aber wegen der notwendigen Einspeicherung der Sicherheitsrangstufen in die Feldtabelle zu einer Erhöhung der Speicherkapazitätsauslastung und zu einer Verlangsamung der Kommunikationsgeschwindigkeit mit den Dateien der Datenbank. Folglich wird sie relativ hohe Kosten verursachen, die die Wirtschaftlichkeit des Datenschutzes in Frage stellen. Darüber hinaus erweist es sich als sinnvoll, die Codeworte innerhalb wechselnder Zeitintervalle zu ändern, um zusätzlichen Schutz zu gewähren.

In der letzten Gruppe der Sicherheitsvorkehrungen bei einer Datenbank sind elektronische Vorrichtungen zu nennen. Diese zeichnen sich dadurch aus, daß sämtliche Zugriffe zur Datenbank registriert und somit nachprüfbar gemacht

¹ Vgl. Schulze, J.H., Datenschutz in der Datenverarbeitung, a.a.O., S. 643.

werden. Sie zeigen an, ob Unbefugte versucht haben, sich Zugriff zur Datenbank zu verschaffen. Das System zählt dabei die Anzahl unberechtigter Zugriffe, um bei Erreichen eines Limits die Verbindung mit diesem Terminal abbrechen zu lassen.¹

¹ Vgl. Schulze, J.H., Datenschutz in der Datenverarbeitung, a.a.O., S. 643; ferner ist auf die Sicherheitsvorkehrungen bei diversen EDV-Systemen hinzuweisen. Vgl. z.B. Pagel, K. / Eberhard, W., Management-Informationssystem MIS /360 - Erste Rationalisierungsstufe im Verwaltungsbereich der IBM Deutschland, Teil I: Anwendungen und Möglichkeiten des MIS/360-Systems, in: IBM Nachrichten, 20. Jg., 1970, S. 463ff, hier S. 468.

4. Softwareentwicklungen für Datenbanken

4.1 Grundprobleme der Datenbank-Software

"Unter Software für Datenbanken und Informationssysteme versteht man die computergerechte Programmierung geeigneter Techniken..."¹ mit dem Ziel, Datenbankverwaltungs- und Datenwiederauffindungsfunktionen selbständig und weitgehend flexibel durch das System vornehmen zu lassen:

"The retrieval function in a data bank will depend on the data organisation. The greater the degree of flexibility is in the data organisation structure the more complicated will the retrieval functions be if the degree of flexibility is to be maintained through the system."²

Das bedeutsamste Problem bei der Entwicklung von Software-Paketen für Datenbanken betrifft einerseits die Implementierung der Datenbanksprache als prozedurale Host-Language andererseits als selbständigen Compiler. Aufgrund der Trennung zwischen diesen Implementierungsarten wird zwischen Operational- und Exekutivsystemen unterschieden.³ Datenbanken, die ausschließlich über Programme zur Datenbankverwaltung verfügen, werden als Operationalsysteme bezeichnet. Bei derartigen Systemen müssen sämtliche Anwendungsprogramme in einer prozeduralen Host-Language abgefaßt werden. Im Gegensatz dazu zeichnen sich Exekutivsysteme dadurch aus, daß sie im Rahmen von Datenbankverwaltungsprogrammen über einen eigenen als selbständigen Compiler implementierten Sprachteil verfügen, der zur Kommunikation mit den Datenbeständen geeignet ist.

1 Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 167.

2 Salomonsson, O., Experiences Of A Pilot Study On Data Banks For Development, in: Journées "Banque De Données", hrsg. von der Association Française pour la Cybernétique Economique et Technique et Institut de Recherche d'Information et d'Automatique, Aix-En-Provence, Juin 1971, S. 15.02ff, hier S. 17.02.

3 Lutz, Th., Generalized Information System (GIS) - Ein Exekutivsystem für Datenbankprobleme, Teil 2: Die GIS-Sprache, in: IBM Nachrichten, 19. Jg., 1969, S. 714ff, hier S. 717.

Für Informationssysteme sind eine Reihe von datenbankspezifischen Software-Paketen entwickelt worden, die hinsichtlich ihrer Anwendbarkeit innerhalb einer gewissen Bandbreite modifizierbar sind.¹ Im folgenden muß auf eine detaillierte Behandlung dieser Programme verzichtet werden, da sie den Rahmen dieser Arbeit sprengen würde. Vielmehr besteht das Ziel der Ausführungen darin, die Grundproblematik dieser Programme anhand eines ausgewählten Beispiels - des Integrated Data Store - zu verdeutlichen.

¹ Es existieren bereits ausgezeichnete Beschreibungen diverser Programme. Beispielhaft sei auf die Veröffentlichung der Gesellschaft für Mathematik und Datenverarbeitung mbH Bonn (GMD), Abteilung für behördliche DV-Systeme, Kurzbeschreibung von Information Storage and Retrieval Systemen, Heft 2, Schloss Birlinghoven, September 1972 hingewiesen. Vgl. ferner den Überblick bei Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 161ff, insbesondere S. 171-185.

4.2 Integrated Data Store (IDS)

Der Integrated Data Store ist eine Datenbank-Organisationstechnik, die die Anwendung spezieller Datenspeicherungs-, -wiederauffindungs- und -fortschreibungsmethoden erlaubt.¹ Er ist hinsichtlich seiner Zugriffsmöglichkeiten eines der effizientesten Systeme und zeichnet sich besonders durch seine Benutzerfreundlichkeit aus.² Durch die zur Anwendung gelangenden Adreßtechniken kommt der IDS der Forderung nach Bildung umfassender logischer Sätze, die unabhängig von ihrer physischen Organisation sind, weitgehend nach. Dieses kommt insbesondere dadurch zum Ausdruck, daß der Benutzer die zwischen den Datenelementen vorliegenden Relationen selbst definieren und im Anwendungszeitpunkt die Zugriffsweise spezifizieren kann. Somit stellt der IDS ein Datenspeicherungs- und -wiederauffindungssystem dar, dessen Datenstruktur durch spezielle Anwendungserfordernisse charakterisiert ist.

Der Integrated Data Store ist ein reines Operationalsystem.³ Folglich verfügt er ausschließlich über Datenverwaltungsprogramme, so daß sämtliche Anwendungsprogramme in konventionellen Programmiersprachen (i.d.R. COBOL oder PL/I) verfaßt werden müssen.⁴

Wesensbestimmendes Merkmal des IDS ist die gekettete Organisation, durch die mehrere Segmente miteinander verbun-

1 Vgl. CODASYL Systems Committee, Feature Analysis Of Generalized Data Base Management Systems, a.a.O., S. 60.

2 Vgl. Schauff, K., Datenbank-Management mit IDS, in: B-GE-Schriften, Heft 4, August 1970, S. 38.

3 Vgl. hierzu Lutz, Th. / Klimesch, H., Die Datenbank im Informationssystem, a.a.O., S. 169, die dem IDS eine Zwischenstellung zwischen Exekutiv- und Operationalsystem mit Tendenz zum Exekutivsystem zusprechen, während für Schauff das IDS ein Operationalsystem ist. Vgl. Schauff, K., Datenbank-Management mit IDS, a.a.O., S. 8ff.

4 Bull-General Electric, IDS Einführung, Ref.-Nr. 18.12.080 D, März 1968, S. 23. IDS arbeitet mit COBOL als Host-Language. Dabei beschreiben COBOL-Sprachelemente die IDS-typischen Speicherungs-, Wiederauffindungs- und Fortschreibungsfunktionen.

den werden. Die Kettenstrukturen des IDS bestehen dabei immer aus einem Hauptsegment (Master- oder Leitsatz) und beliebig vielen Untersegmenten (Detailsätzen).¹ IDS bedient sich dabei der zyklisch geketteten Organisation, bei der das Adreßfeld des ersten logischen Segments auf das logisch erste Segment der Datenbank verweist. Durch eine derartige Adreßtechnik wird die zyklische Kette geschlossen, so daß es aufgrund der zirkularen Anordnung der Segmente unbedeutend ist, bei welchem Segment zuerst mit der Verarbeitung begonnen wird. Die folgende Abbildung verdeutlicht eine derartige Kettenstruktur:²

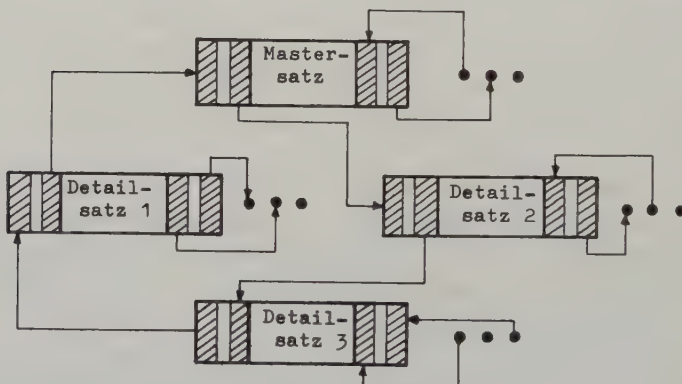


Abb. 8 : Kette im Integrated Data Store

Sowohl Master- als auch Detailsätze stehen im Speicher auf einer Satzadresse und sind über ein oder mehrere Ordnungskriterien identifizierbar. Darüber hinaus enthalten sie vom Benutzer festgelegte Datenelemente in ihren Feldern. Zum Zweck der Datenübertragung zwischen Kernspeicher und

¹ Vgl. zu den Kettenstrukturen Bull-General Electric, IDS Einführung, a.a.O., S. 7ff.

² Vgl. ebenda, S. 11.

externen Speichern mit wahlfreiem Zugriff werden auch im IDS die Sätze zu Blöcken zusammengefaßt.¹

Um Sätze in einem Block zusammenzufassen, enthält jeder Satz zu Beginn drei Datenfelder, und zwar für den Satztyp, den Index und die Satzlänge.² Da IDS sich dadurch auszeichnet, daß verschiedene Sätze miteinander verketten werden können, ist für jedes Glied ein Verbindungselement zu definieren, das auf das nächste Element in der Kette verweist. Die Indizes zwischen den einzelnen Segmenten ermöglichen erst deren adreßtechnische Verkettung. Sie werden im IDS als "chain fields" bezeichnet.³

Bei erstmaliger Abspeicherung eines Segments wird vom IDS eine Transformation dieses Segments vorgenommen, um die Blocknummer der Speicherung bestimmen zu können. Durch eine derartige Vorgehensweise ist es möglich, die relative Positionierung eines Satzes zu bestimmen. Diese erfolgt dadurch, daß dem Satz eine Nummer zugeordnet wird, welche über seine relative Stellung im Verhältnis zu anderen Sätzen in der Kettenstruktur Auskunft gibt. Das Charakteristikum jedes Datensatzes besteht in einer beliebigen Anzahl von Feldern. Er kann mit einer vom Benutzer bestimmten beliebigen Anzahl von Ketten verbunden und braucht nur einmal abgespeichert werden.

Um den Besonderheiten der zyklisch geketteten Organisation im IDS gerecht zu werden, hat es sich eingebürgert, die zyklischen Kettenstrukturen als Ringstrukturen zu bezeichnen. Diese Bezeichnung besagt nichts anderes, als daß sich durch die hierarchische Strukturierung der Sätze im Programm

1 Vgl. Bachman, C.W., Integrated Data Store, in: DPMA Quarterly, January 1965, S. 10ff.

2 Vgl. ebenda, S. 14ff.

3 Vgl. Bachman, C.W. / Williams, S.B., A General Purpose Programming System for Random Access Memories, in: Proceedings of the Fall Joint Computer Conference, Part I, San Francisco, Oktober 1964, S. 411ff, hier S. 413.

Satzstrukturen sowohl intern als auch beliebig untereinander verketteten lassen. In der folgenden Abbildung ist eine derartige Kettung dreier hierarchischer Strukturen mit drei Ringen dargestellt:¹

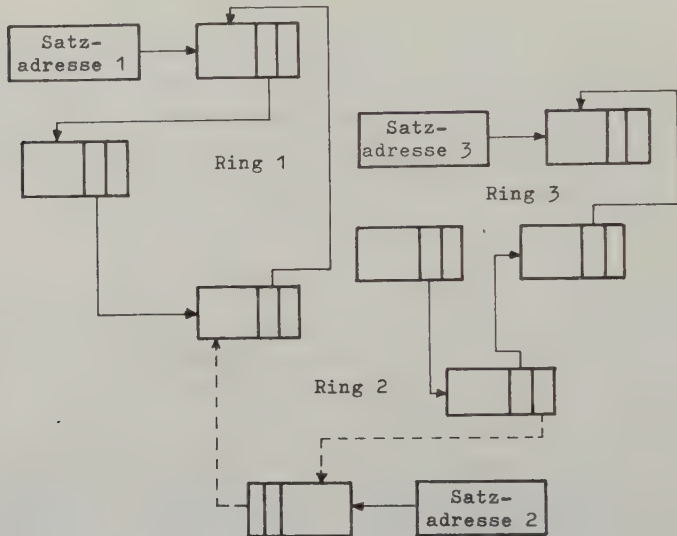


Abb. 9 : Kettung dreier hierarchischer Strukturen mit drei Ringen im Integrated Data Store

Bei einer derartigen Ringstruktur ist immer sichergestellt, daß die Kette geschlossen ist. Jede Kette verfügt dabei über einen Master-Satz. Dieser hat die Aufgabe der Kennzeichnung der Kette. Er kann aber auch für andere Ketten als Master- oder Detailsatz fungieren. Dadurch ist gewährleistet, daß im IDS ein Segment sowohl Master- als auch Detailsatz beliebiger Ketten sein kann.² Die Kettfelder

¹ Vgl. Niedereichholz, J., Betriebswirtschaftslehre und Informatik, in: Wissenschaftsprogramm und Ausbildungsziele der Betriebswirtschaftslehre, Tagungsberichte des Verbandes der Hochschullehrer für Betriebswirtschaft e.V., Bd. I, hrsg. v. G. v. Kortzfleisch, Berlin 1971, S. 225ff, hier S. 234.

² Vgl. die Darstellung der verschiedenen Kombinationsformen bei Schauff, K., Datenbank-Management mit IDS, a.a.O., S. 8.

können dabei auf das nächste Segment, auf das vorangegangene Segment und den Master-Satz der Kette verweisen. Alle Ketten sind zyklisch organisiert, so daß der letzte Detailsatz in der Kette stets auf den Master-Satz zurückverweist. Der Benutzer des IDS hat nun aufgrund von beobachteten empirischen Relationen zwischen den Sätzen zu entscheiden, welche Verkettungen vorgenommen werden sollen. Je mehr empirische Relationen vorhanden sind, um so vielfältiger werden die zur Anwendung gelangenden Kettenstrukturen sein.

Die in einem Master-Satz enthaltenen Datenelemente beziehen sich auf sämtliche Detailsätze, die mit dem Master-Satz verkettet sind. Darüber hinaus ist es dem Benutzer möglich, für jeden Detailsatz innerhalb einer Kette zusätzliche Zeiger zu spezifizieren, um die Hierarchie von Master- und Detailsätzen offenzulegen.

Durch die Ringstruktur des IDS können beliebige hierarchische Strukturen sowohl innerhalb einer Datei als auch zwischen mehreren Dateien verkettet werden.

Die physische Datenorganisation dieser Ketten geschieht in Pages - Seiten fester Länge -, deren Umfang vom Benutzer bei der Dateierrichtung festgelegt wird. Unter einer "Seite fester Länge" wird ein Speicherbereich auf einem Datenträger verstanden, dessen Minimal- und Maximallänge durch einen aufzunehmenden Zeichenvorrat beschränkt ist. Sie besteht aus einem Kennsatz und einem Bereich für Datensätze.² Der gesamte Speicherungsraum des Mediums ist dabei in physisch gleich große Pages eingeteilt, wobei jede durch eine Nummer und einen Veränderungszeiger gekennzeichnet ist. Nummer und Veränderungsanzeiger bilden zusammen

1 Vgl. dazu Spies, P.P., Teilnehmer-Rechensysteme, in: Elektronische Rechanlagen, No. 9, 1967, S. 117ff, hier S. 120f; Bull-General Electric, IDS Einführung, a.a.O., S. 3ff.

2 Zur Technik des Paging im IDS vgl. Bull-General Electric, IDS Einführung, a.a.O., S. 3ff.

den Kennsatz der Page. Die Aufgabe der Nummer besteht darin, die Page wiederauffinden zu können, während der Veränderungsanzeiger darüber Auskunft gibt, ob der Inhalt der in ihr enthaltenen Blöcke während seines Aufenthalts im Kernspeicher der Zentraleinheit verändert wurde oder nicht. Die Individualität und Wiederauffindbarkeit der gespeicherten Datenelemente wird somit durch den Kennsatz und die vom System-Benutzer spezifizierten Datenelemente gewährleistet.

Die Kombination von Page-Nummer und Veränderungszeiger bildet den Segmentschlüssel, der in den Kettenzeigern Anwendung findet. Ist es notwendig, ein Segment auf Veranlassung des Benutzers zu ändern, erfolgt eine Umstrukturierung des Kennsatzes, da es i.d.R. bei Segmentänderungen notwendig sein wird, auch andere Blöcke zu "bewegen", um für das geänderte Segment Speicherplatz zu schaffen. Sollte der maximale Zeichenvorrat in einer Page bei einer vom Benutzer initiierten Segmentänderung ausgeschöpft sein, dann wird vom IDS automatisch eine Speicherung auf dem nächsten freien Speicherplatz in anderen Pages vorgenommen und eine entsprechende Adreßverkettung durchgeführt. Die Werte der Veränderungszeiger müssen dabei entsprechend geändert werden, um die neue Positionierung in der Page angeben zu können. Im IDS erfolgt auch dieser Prozeß durch das System, indem Detail-Sätze in dieselbe Page eingespeichert werden, in der der Master-Satz steht.

Die sich im IDS vollziehenden Prozesse gehen so vor sich, daß zu Beginn der Verarbeitung die Pages, die die zu verarbeitenden Segmente enthalten, in externen Speichern residieren. Diese werden vom System automatisch abgerufen, sobald sie zur Ausführung eines Programms benötigt werden. IDS ist konzeptionell so angelegt, daß die Verweilzeit der im Kernspeicher residierenden Pages so lang wie möglich sein soll, damit die Zugriffsfrequenz zu den externen Speichern möglichst gering gehalten werden kann. Der Ver-

Änderungszeiger erweist sich dabei als notwendig, um dieses Ziel zu erreichen. Wird eine Page zur Ausführung eines Programms in den Kernspeicher übertragen, so bestimmt IDS aufgrund der Veränderungsanzeiger, auf welche Pages in der Vergangenheit nur relativ selten zugegriffen wurde.¹ Diejenige mit der geringsten Zugriffsfrequenz wird dann in den Kernspeicher übertragen. Zeigt der Veränderungszeiger innerhalb eines bestimmten Zeitraums keine Veränderung des Page-Inhalts an, so wird das Segment, das nicht verändert wurde, aus Gründen der Speicherökonomik gelöscht. In Hinblick auf den Zugriff und damit auf die Zugriffszeiten ist die Tatsache bemerkenswert, daß der IDS nicht nur das benötigte Segment in den Kernspeicher überträgt, sondern die gesamte Page, in der dieses Segment enthalten ist. Daraus resultieren erhebliche Verarbeitungsvorteile, insbesondere dann wenn noch weitere Segmente für die Bearbeitung benötigt werden.² Durch diese Vorgehensweise lassen sich zusätzliche Zugriffe zum externen Speicher vermeiden, die anderenfalls unumgänglich wären.

In der folgenden Abbildung ist ein Zugriffsweg beim IDS dargestellt, in der die Pfeile die jeweilige Zugriffsrichtung angeben.³ Das der Abbildung zugrundeliegende Beispiel basiert auf einer für Marketing-Zwecke typischen Problemstellung, bei einer Kundenauftragskette so wohl vom Kundenstammsatz als auch vom Artikelstammsatz auf den Positionssatz zugreifen zu können. So kann es nach der vorliegenden Fragestellung sinnvoll sein, einerseits vom Artikelstammsatz auf den Positionssatz, andererseits vom Kundenstammsatz auf den Positionssatz zuzugreifen.

1 Vgl. Bachman, C.W., Integrated Data Store, a.a.O., S. 16.

2 Vgl. Bull-General Electric, IDS Einführung, a.a.O., S. 6.

3 Vgl. Schauff, K., Datenbank-Management mit IDS, a.a.O., S. 19.

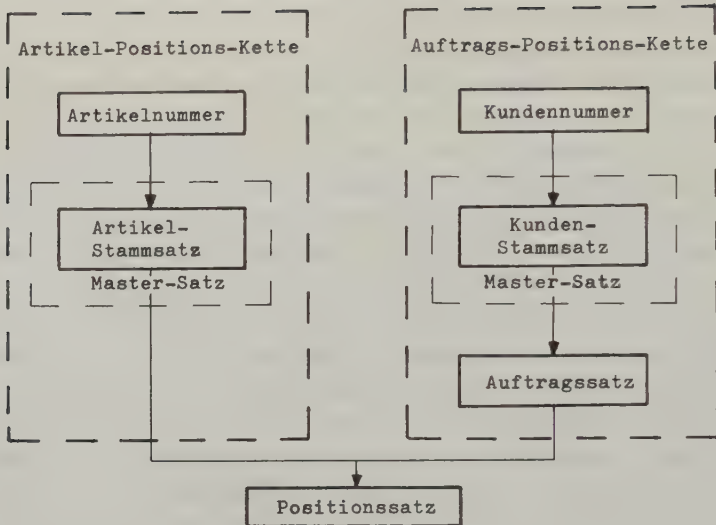


Abb. 10 : Zugriffsweg beim Integrated Data Store

Der Positionssatz ist im Verhältnis zur Artikelpositionskette und zur Auftragspositionskette Detailsatz. Der Artikelstammsatz ist Master-Satz der Artikelpositionskette, der Auftragssatz Master-Satz der Auftragspositionskette. Dieser Auftragssatz ist wiederum in seiner hierarchischen Stellung Detail-Satz der Kundenauftragskette.¹ Beim IDS-Zugriffsweg ist es nun möglich, sowohl vom Master-Satz der Auftragspositionskette als auch vom Master-Satz der Artikelpositionskette auf den Positionssatz zuzugreifen. Daraus resultiert, daß es im IDS gleichgültig ist, auf welchen Master-Satz zuerst zugegriffen wird. Das Beispiel zeigt, daß der Positionssatz immer über zwei verschiedene Zugriffswege erreicht werden kann. Dadurch ist es im IDS er-

¹ Besteht eine Kette aus sehr vielen Detail-Sätzen, kann sie, um die Zugriffszeiten zu den einzelnen Sätzen zu verringern, durch wahlfrei zugreifbare Master-Sätze segmentiert werden. Vgl. Bull-General Electric, IDS Einführung, a.a.O., S. 19.

reicht, daß die physische Datenorganisation für die Formulierung logischer Sätze weitgehend unbedeutend ist.

Zusammenfassend lassen sich die Merkmale des IDS wie folgt angeben:

1. Durch die zirkulare Anordnung der Segmente entsprechend der Ringstruktur des IDS ist es unbedeutend, bei welchem Segment zuerst mit der Verarbeitung begonnen wird.
2. Durch die nur einmalige Speicherung jedes Datensatzes - unabhängig von der Anzahl der Ketten, an denen er partizipiert - wird die benötigte Speicherkapazität erheblich reduziert. Dieses ist der originäre Vorteil, den die Anwendung von IDS bietet.

III. Die Methodenbank im Marketing-Informationssystem

1. Methoden als Komponenten der Methodenbank

1.1 Zum Begriff der Methode

Bei der Erörterung von Problemen der Methodenbank stellt sich die Frage, warum in Anlehnung an MONTGOMERY¹ im Rahmen dieser Arbeit zwischen Modellen und Methoden getrennt wird. Diese Unterscheidung beruht auf der Tatsache, daß die Methodenbank u.a. alle für das Marketing besonders wichtigen statistischen Methoden (z.B. multivariate Datenanalysemethoden) enthält, während die Modellbank die zur Anwendung gelangenden Marketing-Modelle umfaßt.

Die Frage, ob die hier vorgenommene Trennung zwischen Modellen und Methoden sinnvoll ist, enthält nach MEFFERT² ein begriffliches und ein aufbaustrategisches Problem. Wird der Begriff der Methode weit gefaßt - etwa im Sinn eines systematischen Lösungsverfahrens - dann sind die in der Methodenbank eines Marketing-Informationssystems zur Anwendung gelangenden mathematisch-statistischen Verfahren exakte Methoden. Für die exakten Methoden bzw. Modelle besteht eine Prozeßvorschrift, die darüber Auskunft gibt, wie eine gegebene Anfangssituation in einen Endzustand transformiert wird. Bei den inexakten Methoden sind diese Prozeßvorschriften nicht eindeutig interpretierbar.

Darüber hinaus bietet sich die Möglichkeit, statistische Methoden als "taxonomische Modelle" zu bezeichnen. Diese Modelle dienen der Klassifizierung und Systematisierung von Marketing-Phänomenen.³ Das Grundprinzip taxonomischer

1 Vgl. Montgomery, D.B. / Urban, G.L., Marketing Decision-Information Systems, a.a.O., S. 227.

2 Vgl. Meffert, H., Computergestützte Marketing-Informationssysteme und Marketing-Modelle, a.a.O.

3 Vgl. Krautter, J., Marketing-Entscheidungsmodelle, a.a.O., S. 9.

Modelle besteht darin, mehrere Gruppen von Merkmalsträgern mit verschiedenen Variablen in möglichst homogenen Gruppen (wie z.B. bei der Clusteranalyse) zusammenzufassen.

Letztlich besteht der Grund, eine Trennung zwischen Methoden und Modellen vorzunehmen, darin, daß die Methoden der Methodenbank den Input für Marketing-Modelle liefern. Unter Vernachlässigung des begrifflichen Problems, daß exakte Methoden unter dem Begriff des "Modells" subsumiert werden können, soll hier aus Strukturierungs- und Aufbau Gründen eines Marketing-Informationssystems zwischen einer Methoden- und Modellbank unterschieden werden.

1.2 Inhalt und Aufgaben der Methodenbank

Die Methodenbank eines Marketing-Informationssystems enthält Programme von statistischen Verfahren, die zur Auswertung objektiver Daten geeignet sind, von Methoden, die es ermöglichen, Urteilsdaten auszuwerten und von Verfahren, die zum Testen von Hypothesen geeignet sind.¹

Die Notwendigkeit von statistischen Verfahren zur Aufbereitung von Massendaten (z.B. Regressions- und Korrelationsanalyse) ergibt sich aufgrund der Erfahrungstatsache, daß oft eine Diskrepanz zwischen der Darstellungsform dieser Daten und den von verschiedenen Marketing-Funktionsbereichen gewünschten Darstellungsformen besteht. Prinzipiell können Entscheidungsträger jedweden Funktionsbereichs nur dann relevante Daten zur Verfügung gestellt werden, wenn in der Datenbank ein hohes Einzeldaten-Volumen vorhanden ist. Dieses garantiert dafür, daß ein gesuchtes Einzeldatum mit großer Wahrscheinlichkeit aus der Datenbank zu extrahieren ist.

Die Forderung, auf spezielle Einzeldaten zugreifen zu können, gilt in erster Linie für die operative Marketing-Ebene. Während hier der Computer primär dazu herangezogen wird, Daten nach bestimmten Kriterien zu sortieren¹, kommt es dem Marketing-Management insbesondere darauf an, daß homogene Massendaten "verdichtet" werden können. Das Ziel dieser "Verdichtung" besteht darin, aus diesen Daten die für das Marketing-Management wichtigen Trends und Indikatoren abzuleiten. Damit soll das Prinzip des hohen Einzeldatenvolumens in der Datenbank keineswegs in Frage gestellt werden, da ggf. auch das Marketing-Management Einzeldaten benötigt. Tendenziell wird aber der Wunsch nach "verdichteten" Daten dominieren, um daraus entscheidungsrelevante Informationen "zu destillieren".

¹ Vgl. Sheth, J.N., The Multivariate Revolution in Marketing Research, in: Journal of Marketing, Vol. 35, 1971, S. 13ff, hier S. 13.

In diesem Zusammenhang lassen sich zwei Problemkreise unterscheiden, und zwar die Verwendung exakter statistischer Methoden, die auf objektiven Daten basieren und der Einsatz von Verfahren auf Grundlage von Urteilsdaten (wie z.B. Wahrscheinlichkeitsschätzungen und Präferenzfunktionen). Die Tatsache, daß Verfahren in der Methodenbank enthalten sein sollen, die es gestatten, Urteilsdaten auszuwerten, hat seinen Grund darin, daß die Entscheidungsträger im Marketing-Managementsystem oft aufgrund von Daten entscheiden müssen, die zu mehrwertigen Erwartungen führen.

Beispiele für auf Urteilsdaten basierenden Verfahren liefert die statistische Entscheidungstheorie. Sie kann z.B. bei der Auswahl geeigneter Marketing-Strategien in zwei Stufen vorgehen. In der ersten Stufe werden diejenigen Strategien ausgeschieden, die entweder suboptimal sind oder den Marketing-Zielen widersprechen. In dieser Stufe werden alle Marketing-Strategien ausgeschieden, die von anderen dominiert werden. Dann erfolgt die Eliminierung der Strategien mit zu hohen Risiken aufgrund von Risiko-Präferenzfunktionen der Entscheidungsträger. Schließlich wird daran anschließend eine Ausschaltung derjenigen Strategien angestrebt, die keinen gegebenen Mindestgewinn erreichen.

In der zweiten Stufe wird aus den verbleibenden Strategien diejenige ausgewählt, die das Marketing-Management als optimal erachtet. Bei dieser Auswahl, für die es keine eindeutigen Kriterien gibt, können im Rahmen der Methodenbank mathematisch-statistische Entscheidungskriterien verwendet werden. In Abhängigkeit vom jeweiligen Entscheidungskriterium¹ wie z.B. der Maximierung des Erwartungswertes, dem Pessimismuskriterium und dem Hurwicz-Kriterium erweisen sich unterschiedliche Strategien als optimal. Diese Ent-

¹ Zu den Entscheidungskriterien vgl. Heinen, E., Das Ziel-system der Unternehmung, a.a.O., S. 178ff.

scheidungskriterien sind besonders dazu geeignet, auf Erfahrung basierende Schätzungen der Entscheidungsträger zu berücksichtigen. So sind z.B. bei der Maximierung des Erwartungswertes die Eintrittswahrscheinlichkeiten bestimmter Datenkonstellationen zu schätzen; beim Pessimismuskriterium Vorstellungen über die Reihenfolge der schlechtesten Gewinnsituationen zu entwickeln und beim Hurwicz-Kriterium eine Gewichtung der günstigsten und der schlechtesten Gewinnsituation vorzunehmen. Obwohl es mittels dieser auf subjektiven Daten basierenden Verfahren möglich ist, mehrwertige Erwartungen in einwertige zu überführen, bleiben auch hier Erwartungen, die objektiv mehrwertig sind, trotz des Einsatzes von Entscheidungskriterien objektiv mehrwertig.

Schließlich lassen sich im Rahmen der Methodenbank eines Marketing-Informationssystems Verfahren unterscheiden, die zum Testen von Hypothesen geeignet sind. Im Begriff "Hypothesentest" werden unterschwellig gewisse wissenschaftstheoretische Postulate angesprochen. Im folgenden soll gezeigt werden, daß Hypothesentests mit "einfachsten" Methoden zu bewerkstelligen sind. Dieses soll anhand eines Beispiels zur Werbeerfolgskontrolle verdeutlicht werden.

Im Rahmen der Werbeerfolgskontrolle gelangen öfters Experimenttypen zur Anwendung, die sich infolge ihrer einfachen mathematischen Strukturierung aber dennoch großen Aussagefähigkeit für die Aufnahme in die Methodenbank eines Marketing-Informationssystems eignen. Das Marketingproblem kann z.B. darin bestehen, die Wirksamkeit einer Werbekampagne zu testen. Eine repräsentative Verbrauchergruppe, deren Verhalten aus früheren Untersuchungen bekannt ist, wird einer Testanzeige ausgesetzt und nach ihrer Wirkung befragt. Gleichzeitig werden entsprechende Befragungen in einer Kontrollgruppe vorgenommen, die der Testanzeige nicht ausgesetzt war. Es wird die Hypothese aufgestellt, daß Abweichungen bei der zweiten Befragung

auf die fehlende Werbewirkung der Anzeige zurückgeführt werden können. Es stehen vier Experiment-Typen zur Verfügung, um diese Hypothese zu testen.¹

Das Experiment von EBA-Typ (Experimental Group Before and After) arbeitet mit einer Gruppe, die zwei Messungen - vor und nach dem Wirksamwerden des zu testenden Faktors (Werbung) - unterworfen wird. Wird mit x_0 der abhängige Faktor (z.B. pro-Kopf-Verbrauch des umworbenen Produkts) und mit x_1 der abhängige Faktor nach Auswirkung der Werbung bezeichnet, so ist die Faktorwirkung $x_1 - x_0$. Beim Experiment vom EB-CA-Typ (Experimental Group Before - Control Group After) werden zwei repräsentative Gruppen untersucht. Der abhängige Faktor wird vor Auswirkung der Werbung bei der ersten Gruppe gemessen (y_0), während die gleiche Messung nach Auswirkung der Werbung bei der zweiten Gruppe vorgenommen wird (x_1). In diesem Fall beträgt die Faktorwirkung $x_1 - y_0$. Beim Experiment vom EBA-CBA-Typ (Experimental Group Before and After - Control Group Before and After) erfolgt eine Trennung zwischen Personen, die der Werbung ausgesetzt waren (Versuchsgruppe) und die nicht von ihr erreicht worden sind (Kontrollgruppe). Sowohl bei der Versuchs- als auch der Kontrollgruppe wird der abhängige Faktor vor und nach dem Eintritt des Wirkungsfaktors gemessen. In diesem Fall beträgt die Faktorwirkung $(x_1 - x_0) - (y_1 - y_0)$. Schließlich wird beim EA-CA-Typ (Experimental Group After - Control Group After) zwischen einer Versuchs- und Kontrollgruppe unterschieden und eine Messung des Untersuchungsmerkmals nur nach Auswirkung der Werbung vorgenommen. Die Faktorwirkung beträgt hier $x_1 - y_1$.

¹ Vgl. dazu Behrens, K.Ch., Marktforschung, 2. Aufl., Wiesbaden 1966, S. 71ff und Hüttner, M., Grundzüge der Marktforschung, Wiesbaden 1972, S. 133f.

Diese Ausführungen zeigen Überblickhaft, welche Methoden in der Methodenbank eines Marketing-Informationssystems enthalten sein müssen. Ihr Design ist besonders unter dem Gesichtspunkt von Bedeutung, da der Output von Methoden oft in Marketing-Modelle eingeht, der Modelloutput aber nur so gut sein kann wie sein Input. Ausgehend vom Postulat MONTGOMERYs¹, daß die Methodenbank eine Vielzahl multivariater Verfahren, nicht-parametrischer Methoden, Skalierungstechniken und numerische Schätzverfahren enthalten sollte, muß im Rahmen der Erörterung von Methoden ein Schwerpunkt gelegt werden. Infolgedessen werden im folgenden ausschließlich multivariate statistische Verfahren vertiefend behandelt, da von ihnen zu erwarten ist, daß sie sich besonders für Marketing-Fragestellungen eignen.

¹ Vgl. Montgomery, D.B. / Urban, G.L., Marketing Decision-Information Systems, a.a.O., S. 231.

2. Einordnung multivariater Daten-Analysemethoden

"Man spricht von multivariaten statistischen Methoden, wenn ... [bei jedem Objekt] mehrere Variablen gemessen werden, und man alle diese Größen für eine Reihe von .. [Objekten] in der statistischen Analyse gleichzeitig betrachtet und gemeinsam auswertet. ... Es leuchtet unmittelbar ein, daß die gleichzeitige Auswertung mehrerer Variablen im allgemeinen schärfere Aussagen erlauben muß als die Behandlung nur einer Variablen."¹

Die multivariaten statistischen Methoden sind insbesondere dadurch gekennzeichnet, daß ein sehr hoher Rechenaufwand mit ihnen verbunden ist. Erst durch den Einsatz von Computern, die sich durch eine hohe Rechengeschwindigkeit auszeichnen und bei in der Datenbank gespeicherten Massendaten ist es möglich, diese multivariaten statistischen Verfahren sinnvoll einzusetzen. In diesem Zusammenhang spricht SHETH² von der multivariaten Revolution im Marketing Research. Er versteht darunter die Fähigkeit des Computers, große Datenvolumina entsprechend multivariaten Daten-Analysemethoden aufzubereiten und zu verarbeiten. Ein Indikator für die multivariate Revolution im Marketing Research besteht seines Erachtens darin, daß bereits eine große Anzahl von "canned" Computerprogrammen existiert, die im Zusammenhang mit den in der Datenbank gespeicherten Daten zu einer Verbesserung der Information des Marketing-Managements führen werden.³ Die Rolle der elektronischen Datenverarbeitung hat seiner Auffassung nach zu einer starken Verbreitung der multivariaten Datenanalysemethoden im Marketing geführt. Das Fehlen von geeigneten Computerprogrammen war bislang ein Hauptfaktor für die bestehende

1 Überla, K., Faktorenanalyse - Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler, Berlin - Heidelberg - New York 1968, S. 81.

2 Vgl. Sheth, J.N., The Multivariate Revolution in Marketing Research, a.a.O., S. 14.

3 Vgl. ebenda.

Ungleichgewichtigkeit zwischen einem hohen Datenbankvolumen und der extrem unzulänglichen Analyse und Auswertung der in Datenbanken gespeicherten Massendaten.¹ Erst durch die Symbiose von hoher Rechengeschwindigkeit des Computers, vorhandenen Methoden der multivariaten Datenanalyse und gespeicherten Massendaten kann dem Postulat des managerorientierten Ansatzes entsprochen werden, daß durch eine Verdichtung der dem Marketing-Management zugehenden Daten bessere Entscheidungen gefällt werden können.

In diesem Zusammenhang scheint es sinnvoll zu sein, diejenigen multivariaten Daten-Analysemethoden zu untersuchen, die in der Methodenbank eines Marketing-Informationssystems wesentlich sind. Für die Anwendung multivariater statistischer Methoden sind zwei Gründe maßgeblich:² Der erste Grund besteht darin, daß bereits in der Psychologie eine Reihe von multivariaten statistischen Daten-Analysemethoden existieren, die bei Marketing-Problemstellungen angewendet werden können. Der zweite Gesichtspunkt für die Anwendung multivariater Methoden ist darin begründet, daß uni- und bivariate Verfahren der Datenanalyse sich als wenig geeignet erwiesen haben, komplexe Marketing-Phänomene zu untersuchen. Dieser Gedanke konkretisiert sich in folgender Feststellung: "Für Zwecke der Marktforschung oder irgendeines anderen angewandten Gebiets sind die meisten Instrumente multivariat oder sollten es zumindest sein. Es drängt sich einem die Schlußfolgerung auf, daß die Behandlung eines Marketing-Problems dann überflüssig ist, wenn es nicht als multivariates Problem behandelt wird."³

¹ Vgl. dazu die Ausführungen bei Cooley, W.W. / Lohnes, P.R., *Multivariate Data Analysis*, New York 1971, insbesondere die dort dargestellten Programme zur multivariaten Datenanalyse.

² Vgl. Sheth, J.N., *Multivariate Analysis in Marketing*, in: *Journal of Advertising Research*, Vol. 10, 1970, S. 29ff, hier S. 29.

³ Gatty, R., *Multivariate Analysis for Marketing Research - An Evaluation*, in: *Applied Statistics*, Vol. 15, 1960, S. 157ff, hier S. 157. (Übersetzung vom Verfasser)

Für die Einordnung¹ der bekanntesten multivariaten Daten-Analysemethoden stehen zwei Klassifizierungskriterien zur Verfügung. Das erste Klassifizierungskriterium knüpft daran an, ob die zu untersuchenden Variablen voneinander unabhängig sind oder nicht. Es können hier Verfahren unterschieden werden, die die Untersuchung von Abhängigkeiten, Zusammenhängen und Klassifikationen bzw. Typisierungen ermöglichen. Das zweite Klassifizierungskriterium betrifft die Art und Weise, wie die in die Analyse eingehenden Daten skaliert sind. Es lassen sich metrisch, ordinal und nominal skalierte Daten unterscheiden. Je nach Skalierung der Inputdaten bestehen bei den multivariaten statistischen Daten-Analysemethoden spezielle statistische Probleme.

Bei metrischer Skalierung bestehen quantitative Merkmalsausprägungen, die es gestatten, die Daten intervall- oder quotientenweise zu skalieren. Bei nicht metrischer Skalierung, die bei qualitativen Daten zur Anwendung gelangt, können die Daten entweder nominal oder ordinal skaliert werden.

Es läßt sich folgende Einordnung multivariater statistischer Daten-Analysemethoden, bei denen mehrere Merkmals-träger (Objekte) mit mehreren Variablen untersucht werden, vornehmen:

¹ Vgl. dazu Sheth, J.N., The Multivariate Revolution in Marketing Research, a.a.O., S. 15; Überla, K., Faktorenanalyse, a.a.O., S. 82; Jahn, W. / Vahle, H., Die Faktorenanalyse und ihre Anwendung, Berlin 1970, S. 22.

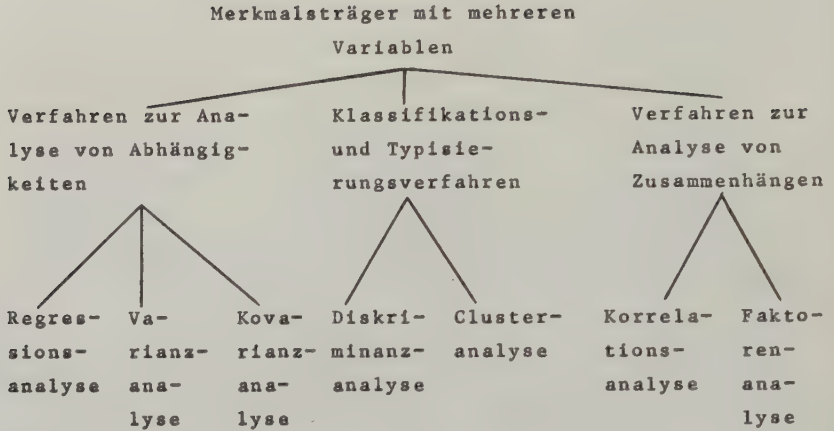


Abb.11 : Systematik multivariater Daten-Analyse-
methoden

Die Datenanalyseverfahren zur Untersuchung von Abhängigkeiten unterstellen Monokausalität, d.h., daß nur eine zu erklärende Variable von mehreren Variablen abhängt. Ein Kausalzusammenhang stellt sich hier als Funktion zwischen einer zu erklärenden und mehreren erklärenden Variablen dar.¹ Eine Regressionsanalyse ist dann anwendbar, wenn die Variablen metrisch skaliert sind. Eine Varianzanalyse setzt ebenfalls Monokausalität² voraus und ist dann anwendbar, wenn die Variablen metrisch oder nominal skaliert sind. Bei der Kovariananalyse sind metrische, nominale und ordinale Skalierungen möglich.

Im Rahmen dieser Arbeit werden ausschließlich Probleme multivariater Daten-analysemethoden bei metrischer Skalie-

¹ Vgl. dazu den Punkt V./3.2 dieser Arbeit

² Zum Zusammenhang zwischen Kausalität und Korrelations- und Regressionsanalyse vgl. z.B. Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse - Ein Leitfaden für Ökonomen, Berlin 1966, S. 11ff.

rung erörtert. Nicht-metrische Skalierungen führen zu statistischen Spezialproblemen, die den Rahmen dieser Untersuchung sprengen würden. Die multivariaten Datenanalysemethoden müssen bestimmten Verteilungsannahmen der in ihnen enthaltenen Variablen und bestimmten Modellannahmen genügen. Diese Annahmen werden bei der Diskussion der einzelnen Methoden untersucht.

3. Verfahren zur Analyse von Abhängigkeiten

3.1 Regressionsanalyse

Aufgabe der Regressionsanalyse ist die Untersuchung monokausaler Abhängigkeiten zwischen metrisch skalierten erklärenden Zufallsvariablen und einer metrisch skalierten zu erklärenden Zufallsvariablen.¹ "Diese einseitige stochastische Abhängigkeit wird mit Hilfe einer Funktion sichtbar gemacht, die im Unterschied zu den streng mathematischen Funktionen als Regressionsfunktion bezeichnet wird."² Mittels zweier Denkmodelle kann die Abhängigkeit zwischen der zu erklärenden Zufallsvariable und den erklärenden Zufallsvariablen offengelegt werden.

Im ersten Denkmodell³ wird angenommen, daß zwischen der zu erklärenden Zufallsvariable und den erklärenden Zufallsvariablen eine deterministische Abhängigkeit besteht, die durch stochastische Schwankungen überlagert wird. Der funktionale Zusammenhang kann nach Eliminierung der Störvariablen untersucht werden.

Im zweiten Denkmodell⁴ wird davon ausgegangen, daß die Abhängigkeit zwischen zu erklärender Variable und den erklärenden Variablen stochastischer Natur ist. In diesem Fall ist eine funktionale Umkehrung nicht mehr möglich. Zwar können aus sachlogischen Gründen Umkehrungen zwischen zu erklärender Variable und erklärenden Variablen möglich sein, die ermittelten Regressionsfunktionen sind aber nicht mehr umkehrbar. Beim Vorliegen von stochastischen Beziehungen zwischen den Variablen ist die Festlegung der erklärenden Variablen und der zu erklärenden Variable untersuchungszweckabhängig. Daraus folgt, daß die als Regressionsfunktionen unterstellten Funktionstypen nicht a-priori

1 Vgl. Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 25.

2 Ebenda, S. 25f.

3 Vgl. ebenda, S. 31ff.

4 Vgl. ebenda, S. 26.

gegeben sind, sondern aufgrund der Analyse von Streudiagrammen nach heuristischen Gesichtspunkten festgelegt werden müssen.¹

Die Regressionsanalyse soll im einzelnen folgende Aufgaben lösen:²

1. Ermittlung der funktionalen Form und des Verlaufs des Zusammenhangs. Mit heuristischen Grundüberlegungen möglicher Funktionsformen und den damit zusammenhängenden Verläufen der Funktionen ist bei Vornahme der Regressionsanalyse einzusetzen.
2. Feststellung der Wirkungen, die eine oder mehrere Erscheinungen auf die Folgeerscheinung haben. Bei dieser Aufgabenstellung geht es primär darum, die Faktoren zu berücksichtigen, die einen Zusammenhang am stärksten bestimmen.
3. Prognose unbekannter Werte der zu erklärenden Variable bei bekannten oder prognostizierten erklärenden Variablen.

Diese Aufgaben können nur dann gelöst werden, wenn es gelingt, Kriterien für die Güte der stochastischen Abhängigkeiten zu finden. Ein Gütekriterium ist durch die Bedingung gegeben, daß die Summe aller Abweichungen der effektiv beobachteten Werte von denen durch die Regressionsfunktion erklärten Werte (Summe der Residuen) Null ist.

¹ Vgl. Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 26.

² Vgl. ebenda, S. 31ff.

Die Regressionsanalyse arbeitet mit folgenden Annahmen:¹

1. Der Zusammenhang zwischen zu erklärender Variable und den erklärenden Variablen wird i.d.R. als linear angenommen oder linearisiert.
2. Die Fehlerverteilung ist konstant oder homoskedastisch. Daraus folgt, daß die Streuung zwischen den tatsächlich beobachteten und den mittels der Regressionsfunktion erklärten Werten bei sämtlichen erklärenden Variablen konstant ist. Weiterhin wird angenommen, daß die Fehler normalverteilt sind.
3. Es wird angenommen, daß keine Multikollinearität² (die erklärenden Variablen sind voneinander abhängig) zwischen den erklärenden Variablen besteht.

Die Bedingung, daß die Summe der Residuen Null ist, ist nur dann gegeben, wenn sich positive und negative Abweichungen der Beobachtungswerte von der Regressionsfunktion ausgleichen. Unter dieser Bedingung sind aber unendlich viele Regressionsfunktionen möglich, "nämlich alle diejenigen, die durch einen Punkt gehen, der durch die Mittelwerte der Veränderlichen ... bestimmt ist."³ Es ist somit durch eine zusätzliche Bedingung sicherzustellen, daß die zu ermittelnde Regressionsfunktion eindeutig ist. Eindeutigkeit ist dann gegeben, wenn die Quadratsumme der Residuen ein Minimum wird. Diese zusätzliche Bedingung umfaßt die anfangs aufgestellte Forderung und sorgt dafür, daß es nur eine Regressionsfunktion gibt, die beiden Bedingungen zugleich genügt.

1 Vgl. Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, Englewood, Cliffs, N.J. 1967, S. 67.

2 Zu den Problemen der Multikollinearität vgl. Förster, E. / Egernmayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 210ff.

3 Ebenda, S. 57.

Die Ableitung der Normalgleichungen der Regressionsfunktion erfolgt, indem in die Abweichungsquadratsumme die durch die Regressionsfunktion erklärten Werte eingesetzt werden und diese Gleichung nach den Koeffizienten der Regressionsfunktion differenziert und gleich Null gesetzt wird. Mit

$$(1) \quad y_i = a + bx_i + cz_i$$

sei eine lineare Mehrfachregressionsfunktion bezeichnet. Es gilt, daß die Summe der Abweichungsquadrate ein Minimum werden soll:

$$(2) \quad S = \sum_i (y_i - a - bx_i - cz_i)^2 \longrightarrow \text{Min.}!$$

Durch die Differentiation von (2) nach den zu bestimmenden Koeffizienten a, b und c ergibt sich

$$(3) \quad \frac{dS}{da} = -2 \sum_i (y_i - a - bx_i - cz_i) = 0$$

$$\frac{dS}{db} = -2 \sum_i x_i (y_i - a - bx_i - cz_i) = 0$$

$$(5) \quad \frac{dS}{dc} = -2 \sum_i z_i (y_i - a - bx_i - cz_i) = 0$$

Werden in den Gleichungen (3) bis (5) die Summationen über jedem Glied vorgenommen, ergeben sich folgende Normalgleichungen:¹

$$(6) \quad \sum_i y_i = n a + b \sum_i x_i + c \sum_i z_i$$

$$(7) \quad \sum_i y_i x_i = a \sum_i x_i + b \sum_i x_i^2 + c \sum_i z_i x_i$$

$$(8) \quad \sum_i y_i z_i = a \sum_i z_i + b \sum_i x_i z_i + c \sum_i z_i^2$$

¹ Zum allgemeinen Bildungsgesetz der Normalgleichungen der Regressionsfunktion, das die Differentiation erspart, vgl. Fletcher, A. / Clarke, C., Mathematische Hilfsmittel der Unternehmensführung, München 1966, S. 128ff.

Mittels Determinantenrechnung lassen sich die Koeffizienten dieser Normalgleichungen bestimmen.¹

Es ergibt sich eine erhebliche Vereinfachung bei der Berechnung der Regressionskoeffizienten, wenn (6) durch die Anzahl der Beobachtungswerte n dividiert wird:

$$(7) \quad \bar{y} = a + b \bar{x} + c \bar{z}.$$

In (7) stellen $\bar{y}$, $\bar{x}$ und $\bar{z}$ die Mittelwerte der Variablen dar. Wird aus (7) a isoliert und in die Regressionsfunktion $y'_i = a + bx_i + cz_i$ eingesetzt, erhält man:

$$(8) \quad y'_i - \bar{y} = b (x_i - \bar{x}) + c (z_i - \bar{z}).$$

Der nächste sich nach der Berechnung der Koeffizienten der Regressionsfunktion anschließende Schritt besteht in der Untersuchung der Güte dieser Regressionsfunktion.

Es ergibt sich bei der Regressionsanalyse dadurch ein Fehler, daß die beobachteten Werte der erklärenden Variablen um die durch die Regressionsfunktion erklärten Werte streuen. Je geringer diese Streuung ist, um so besser sind die durch die Regressionsfunktion erklärten Werte. Der Fehler wird als Standardfehler der Regressionsschätzung $s_{y/xz}$ ² bezeichnet. Er errechnet sich nach der Formel

$$(9) \quad s_{y/xz} = \sqrt{\frac{\sum_i [y_i - f(x_i; z_i)]^2}{n}}.$$

¹ Vgl. dazu Adam, A. / Fersch, F. / Klamecker, A. et al., Anwendung der Matrizenrechnung auf wirtschaftliche und statistische Probleme, 2. Aufl., Würzburg 1963 und Förster, E. / Egermayer, F., Regressions- und Korrelationsanalyse, a.a.O., S. 83.

² Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, 2. Aufl., New York usw. 1966, S. 125f.

In (9) bedeutet y_i der tatsächlich beobachtete Wert der zu erklärenden Zufallsvariablen, n die Anzahl der Beobachtungswert-Kombinationen und $y_i' = f(x_i; z_i)$ der durch die Regressionsfunktion erklärte Wert. Die Differenz zwischen dem tatsächlich beobachteten und dem durch die Regressionsfunktion erklärten Wert ist ein Residuum oder Schätzfehler. Die Quadratsumme der Residuen ist unter der Voraussetzung einer kleinste-Quadrate-Regressionsfunktion ein Minimum (Zähler des Standardfehlers der Regressions-schätzung).

Dieser Standardfehler der Regressionsschätzung läßt sich im Fall der linearen Zeifachregression berechnen, indem in (9) für $f(x_i; z_i)$ die Regressionsfunktion eingesetzt wird und die Wurzel durch Quadrieren eliminiert wird:

$$(10) \quad s_{y/xz}^2 = \frac{1}{n} \sum_i [y_i - a - bx_i - cz_i]^2$$

$$= \frac{1}{n} (\sum_i y_i^2 - a \sum_i y_i - b \sum_i x_i y_i - c \sum_i z_i y_i) .$$

Werden die Beobachtungswert-Kombinationen, die der Berechnung der Regressionskoeffizienten zugrundeliegen, als Zufallsstichproben interpretiert, die voneinander unabhängig und deren Grundgesamtheiten normalverteilt sind, so kann aus der Stichprobe die Standardabweichung der Regressionsschätzung berechnet werden. Bei einer unendlich großen Grundgesamtheit berechnet sich der unverzerzte Schätzwert für die Standardabweichung der Regressionsschätzung $\hat{\sigma}_{y/xz}$ in der Grundgesamtheit nach der Formel¹

$$(11) \quad \hat{\sigma}_{y/xz} = s_{y/xz} \sqrt{\frac{n}{n-3}}$$

¹ Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 184ff.

mit drei Freiheitsgraden $n - 3$, da durch die Koeffizienten a , b und c drei Freiheitsgrade verloren gehen.¹ Bei einer Grundgesamtheit, die kleiner als 30 ist, entspricht die Verteilung der Stichprobenregressionskoeffizienten einer Student-t-Verteilung, wenn gilt, daß die abhängigen Variablen annähernd normalverteilt sind.

Werden aus normalverteilten Grundgesamtheiten Stichproben gezogen und die Regressionskoeffizienten berechnet, so ergeben sich Regressionsfunktionen, die unterschiedliche Steigungen aufweisen. Von den Koeffizienten b und c der Regressionsfunktion wird angenommen, daß sie um die Steigungsmittelwerte normalverteilt sind. Durch Bestimmung der Standardabweichung des Regressionskoeffizienten b

$$(12) \quad s_b = \frac{s_{y/xz}}{s_x \sqrt{n}}$$

und der Standardabweichung des Regressionskoeffizienten c

$$(13) \quad s_c = \frac{s_{y/xz}}{s_x \sqrt{n}}$$

und durch Einsetzen verzerrungsfreier Schätzwerte für $s_{y/xz}$, s_x und s_y in (12) und (13) lassen sich die Standardfehler der Regressionskoeffizienten bestimmen.

¹ Vgl. Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 199f.

3.2 Varianzanalyse

Die Varianzanalyse ist eine Methode, um die bei experimentellen Daten beobachtete Varianz in verschiedene Teile aufzuspalten. Jeder Teil wird einem bekannten Faktor zugewiesen, um zu untersuchen, ob ein Zusammenhang zwischen Beobachtungswerten und einer Einflußgrößenkategorie (Faktor) oder mehreren Einflußgrößenkategorien besteht.¹ Der erste Fall wird als einfache Klassifikation², der zweite Fall als Mehrfachklassifikation³ bezeichnet. Der Faktor ist nominal skaliert, während die Beobachtungswerte metrisch skaliert sind.

Ein Beispiel soll die Anwendung der Varianzanalyse im Fall der Einfachklassifikation näher verdeutlichen. Eine Unternehmung bedient sich bei der Einführungswerbung für ein neues Produkt drei verschiedener Werbemedien, und zwar des Fernsehens (TV), der Zeitschriftenanzeige (Z) und der nationalen Plakatierung (P). In jedem Medium wird der Konsument aufgefordert, sich unter Nennung des Mediums, in dem er die Einführungswerbung gesehen hat, eine Probe zuschicken zu lassen. Es werden zu drei verschiedenen Zeitpunkten (1 = 1,2,3) die Anfragen der Kunden nach Produktproben erfaßt und dem jeweiligen Medium zugeordnet. Die Untersuchungsergebnisse sind in folgender Tabelle zusammengefaßt:

-
- 1 Vgl. Jones, L.V., Analysis of Variance in its Multivariate Developments, in: Cattell, R.B., Hrsg., Handbook of Multivariate Experimental Psychology, Chicago 1966, S. 244ff, hier S. 245.
 - 2 Vgl. Roßner, R., Varianzanalyse: Einfachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden I: Grundlagen und Versuchsplanung, hrsg. v. E. Walter, Berlin - Heidelberg - New York 1970, S. 217ff.
 - 3 Vgl. derselbe, Zweifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 230ff.

Beob- achtungs- zeitpunkt \ Medium	TV	Z	P	$x_{1.}$
	1	2	3	
1	$x_{11} = 10$	$x_{12} = 8$	$x_{13} = 11$	$x_{1.} = 29$
2	$x_{21} = 7$	$x_{22} = 9$	$x_{23} = 0$	$x_{2.} = 16$
3	$x_{31} = 7$	$x_{32} = 10$	$x_{33} = 1$	$x_{3.} = 18$
$x_{.j}$	$x_{.1} = 24$	$x_{.2} = 27$	$x_{.3} = 12$	$x_{..} = 63$
$x_{.j}$	$x_{.1} = 8$	$x_{.2} = 9$	$x_{.3} = 4$	$x_{..} = 7$
$x_{..} - x_{.j}$	$x_{..} - x_{.1} = -1$	$x_{..} - x_{.2} = -2$	$x_{..} - x_{.3} = 3$	

Tab. 1 : Ausgangsdaten der einfachen Varianzanalyse

Die materielle Marketing-Fragestellung besteht darin, ob ein Zusammenhang zwischen der Werbewirksamkeit der Medien und der Anzahl der Anfragen nach Produktproben besteht.

Diese materielle Fragestellung ist mit der statistischen identisch, ob die Mittelwerte $x_{.j}$ der als Stichprobe zu interpretierenden Beobachtungswerte x_{ij} homogen sind oder nicht.¹ Bei Homogenität der Stichprobenmittel kann unterstellt werden, daß sie aus einer gemeinsamen Grundgesamtheit stammen, während bei Heterogenität der Stichprobenmittel angenommen werden muß, daß die Beobachtungswerte aus verschiedenen Grundgesamtheiten stammen. Das Grundproblem der Varianzanalyse besteht nun darin, ob statistische Signifikanz zwischen Unterschieden in durchschnittlichen Wirkungen besteht, die durch kontrollierte Variablen hervorgerufen werden, und zwar unter Berücksichtigung zufallsbedingter Wirkungseinflüsse unkontrollierter Variablen.

¹ Vgl. Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 33; Überla, K., Faktorenanalyse, a.a.O., S. 84.

Es läßt sich ausgehend vom angeführten Beispiel die Nullhypothese aufstellen, daß zwischen den Faktoren (der Menge aller kontrollierten Variablen) kein Unterschied besteht. Von den unbekannten Mittelwerten der Grundgesamtheiten wird also angenommen, daß $\mu_1 = \mu_2 = \mu_3 = \mu$ ist. Dieser Nullhypothese ist eine Gegenhypothese zugeordnet, daß von den Mittelwerten der j Grundgesamtheiten mindestens zwei Mittelwerte unterschiedlich sind.

Aufgabe der Varianzanalyse ist es, die aus verschiedenen Quellen resultierenden Streuungen dahingehend zu untersuchen, ob ein bestimmter Varianzanteil größer oder kleiner gleich ist als die entsprechende Erwartung bei Geltung der Nullhypothese. In ihrer einfachsten Form ist die Varianzanalyse ein Signifikanztest von Unterschieden zwischen unterschiedlichen Stichprobenmittelwerten. Bei nur zwei Mittelwerten führt dieser Ansatz zu dem gleichen Ergebnis wie ein Signifikanztest zwischen Mittelwerten zweier voneinander unabhängiger Zufallsstichproben.

Aufgrund dieser Zielsetzung muß die Varianzanalyse bestimmten statistischen Annahmen genügen, die sich aufgrund des Zufallscharakters der als Stichprobenwerte interpretierten Beobachtungsdaten ergeben:¹

1. Es wird angenommen, daß die Elemente der Grundgesamtheiten, aus denen die Stichproben gezogen werden, normalverteilt sind.
2. Weiterhin wird unterstellt, daß die Varianzen der Grundgesamtheiten homogen sind.
3. Es wird von additiven Effekten eines Faktors (Einfachklassifikation) oder mehrerer Faktoren (Mehrfachklassifikation) auf die Gesamtvarianz ausgegangen.

¹ Vgl. Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 38.

4. Die Fehler der Beobachtungswerte sind normalverteilt.
5. Die Beobachtungswerte sind voneinander unabhängig; jeder Beobachtungswert wird als unabhängiger Stichprobenwert aus normalverteilten Fehlerverteilungen interpretiert.

Bei der Varianzanalyse im Fall der Einfachklassifikation wird die Additivität der Effekte eines Faktors durch folgende Modellgleichung zum Ausdruck gebracht:¹

Def.
(1) $x_{ij} = \mu + \beta_j + e_{ij}$

In dieser Modellgleichung bedeutet x_{ij} der i -te Beobachtungswert ($i = 1, 2, \dots, m$) der j -ten ($j = 1, 2, \dots, n$) Stichprobe. Durch $\mu_1 = \mu_2 = \mu_3 = \dots = \mu$ wird Nullhypothese zum Ausdruck gebracht, daß die Grundgesamtheit-Mittelwerte gleich sind. Mit β_j wird der Effekt j in der Grundgesamtheit bezeichnet. Er ist die Differenz zwischen den Grundgesamtheit-Mittelwerten, wenn β_j ungleich Null ist. Aus der Gleichheit der zu schätzenden Mittelwerte in der Grundgesamtheit folgt, daß $\mu - \mu_j = 0$ ist, so daß definitionsgemäß

(2) $\sum_j \beta_j = 0$ ist.

Mit e_{ij} wird der Zufallsfehler beim i -ten Beobachtungswert in der j -ten Stichprobe bezeichnet. Da Normalverteilung des Zufallsfehlers angenommen wird, gelten die Annahmen:

(3) $\mu(e_{ij}) = 0$

(4) $\text{Var}(e_{ij}) = \sigma_e^2$.

¹ Die Darstellung des Modells der einfachen Varianzanalyse erfolgt in Anlehnung an Roßner, R., Varianzanalyse: Einfachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 217ff.

Um nach der Modellgleichung (1) die Varianzanalyse vornehmen zu können, ist es notwendig, die unbekannten Parameter μ , β_j und e_{ij} zu schätzen. Die Varianzanalyse geht dabei in vier Schritten vor:¹

Im ersten Schritt wird die Gesamtsumme der Abweichungsquadrate der Beobachtungswerte in zwei additive Komponenten aufgespalten, und zwar in die Summe der Abweichungsquadrate der Beobachtungswerte vom jeweiligen Stichprobenmittel und in die Summe der Abweichungsquadrate der Stichprobenmittel vom Gesamtmittelwert, multipliziert mit der Anzahl der Stichprobenwerte. Wird mit x_{ij} der Beobachtungswert i in der j -ten Stichprobe, mit $x_{.j}$ der Stichprobenmittelwert und mit $x_{..}$ der Gesamtmittelwert der j -ten Stichprobe bezeichnet, so gilt folgende Identitätsgleichung:²

$$(5) \sum_{ij} (x_{ij} - x_{..})^2 = \sum_{ij} (x_{ij} - x_{.j})^2 + \sum_j (x_{.j} - x_{..})^2$$

Durch Quadrieren von (5) ergibt sich

$$(6) \sum_{ij} (x_{ij} - x_{..})^2 = \sum_{ij} (x_{ij} - x_{.j})^2 + 2 \sum_{ij} (x_{ij} - x_{.j})(x_{.j} - x_{..}) + \sum_j (x_{.j} - x_{..})^2$$

Da $\sum_{ij} (x_{ij} - x_{.j})(x_{.j} - x_{..}) = n \sum_j (x_{.j} - x_{..})(x_{.j} - x_{..})$ ist, kann (6) in

$$(7) \sum_{ij} (x_{ij} - x_{..})^2 = \sum_{ij} (x_{ij} - x_{.j})^2 + n \sum_j (x_{.j} - x_{..})^2$$

umgeformt werden.

Der Ausdruck (7) besagt, daß die Gesamtsumme der Abweichungsquadrate (linker Term von (7)) in eine Intragruppen-

¹ Vgl. dazu Ferguson, G.A. Statistical Analysis in Psychology and Education, a.a.O., S. 291.

² Vgl. ebenda, S. 285f.

Abweichungsquadratsumme (erster rechter Term von (7)) und eine Intergruppen-Abweichungsquadratsumme multipliziert mit der Anzahl der Beobachtungswerte (zweiter rechter Term von (7)) aufgespalten werden kann.

Der zweite Schritt besteht in der Bestimmung des Freiheitsgrades obiger Abweichungsquadratsummen. Die Gesamtzahl der Beobachtungswerte beträgt $n \cdot m = N$, bei $i = 1, 2, \dots, m$ und $j = 1, 2, \dots, n$. Ein Freiheitsgrad geht dadurch verloren, daß die Abweichungsquadrate vom Gesamtmittelwert $\bar{x}_{..}$ bestimmt werden. Die Gesamtsumme der Abweichungsquadrate hat somit einen Freiheitsgrad von $N-1$. Die Intragruppen-Abweichungsquadratsumme hat folglich einen Freiheitsgrad von $N-n$, so daß der Freiheitsgrad der Intergruppen-Abweichungsquadratsumme aus der Identitätsgleichung¹

$$(8) \quad N-1 = (N-n) + (n-1)$$

bestimmt werden kann. Die Abweichungsquadratsummen werden dann durch die ihnen zugehörigen Freiheitsgrade dividiert, um die Varianzen zu errechnen. Während die Abweichungsquadratsummen additiv sind, gilt dieses nicht mehr für die errechneten Varianzen, da die Freiheitsgrade der Abweichungsquadratsummen unterschiedlich sind. Die Gesamtvarianz beträgt somit

$$(9) \quad s_{\text{gesamt}}^2 = \frac{\sum \sum (x_{ij} - \bar{x}_{..})^2}{N-1},$$

die Intragruppenvarianz

$$(10) \quad s_{\text{intra}}^2 = \frac{\sum \sum (x_{ij} - \bar{x}_{.j})^2}{N-n}$$

und die Intergruppenvarianz

$$(11) \quad s_{\text{inter}}^2 = \frac{n \cdot \sum (\bar{x}_{.j} - \bar{x}_{..})^2}{n-1}.$$

¹ Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 286.

Diese drei Varianzen stellen Schätzwerte für die unbekannten Varianzen der Grundgesamtheit dar.

Der nächste zu vollziehende Schritt besteht in der Bestimmung des F-Quotienten aus s_{inter}^2 und s_{intra}^2 :¹

$$(12) \quad F^{\text{emp.}} = \frac{s_{\text{inter}}^2}{s_{\text{intra}}^2} .$$

Ist $F^{\text{emp.}}$ größer als $F^{\text{theor.}}$ bei vorgegebener Irrtumswahrscheinlichkeit von $F^{\text{theor.}}$, so wird die Nullhypothese von der Gleichheit der Mittelwerte μ_j in der Grundgesamtheit verworfen und die Alternativhypothese angenommen.

Aufgrund des einleitend angeführten Beispiels ergibt sich eine Intergruppen-Abweichungsquadratsumme von $3|(8-7)^2 + (9-7)^2 + (4-7)^2| = 42$. Der ihr zugehörige Freiheitsgrad $f_1 = n-1$ beträgt 2. Die Intergruppenvarianz beträgt somit 21. Die Intragruppen-Abweichungsquadratsumme lautet $(10-8)^2 + (7-8)^2 + (7-8)^2 + (8-9)^2 + (9-9)^2 + (10-9)^2 + (11-4)^2 + (0-4)^2 + (1-4)^2 = 82$. Da ihr zugehöriger Freiheitsgrad $f_2 = N-n = 6$ ist, beträgt die Intragruppenvarianz somit 13,33. Da $F_{0,95;2,6}^{\text{theor.}} = 5,143$ größer als $F^{\text{emp.}} = 1,57$ ist, muß die Nullhypothese angenommen werden, daß die drei Werbemedien hinsichtlich ihrer Eignung für die Einführungswerbung gleich gut geeignet sind.

Die zweifache Klassifikation zeichnet sich dadurch aus, daß zwei Faktoren mit ihren verschiedenen nominal skalierten Ausprägungen untersucht werden. Wiederum soll ein Beispiel die Fragestellung bei der Zweifachklassifikation verdeutlichen. Ein Unternehmen will drei Waschmittel ($j = 1, 2, 3$) dahingehend überprüfen, von welchen Hausfrauen-Altersklas-

¹ Vgl. Jones, L.V., Analysis of Variance in its Multivariate Developments, in: Handbook of Multivariate Experimental Psychology, a.a.O., S. 248.

sen diese überwiegend gekauft werden. In jeweils zwei Versuchen testet die Unternehmung den Waschmittelabsatz in zwei Altersklassen i_1 (von 20 bis unter 40 Jahren) und i_2 (von 40 Jahren bis unter 60 Jahren). Die Untersuchung hat folgende Werte ergeben:

Hausfrauen- Altersklas- se i		Waschmittel j			$X_{i..}$ $x_{i..}$
		$j = 1$	$j = 2$	$j = 3$	
Al- ters- klas- se 1	Versuch 1	$x_{111} = 6$	$x_{121} = 18$	$x_{131} = 36$	
	Versuch 2	$x_{112} = 12$	$x_{122} = 12$	$x_{132} = 42$	
$X_{1j.}$		$X_{11.} = 18$	$X_{12.} = 30$	$X_{13.} = 78$	$X_{1..} = 126$
$x_{1j.}$		$x_{11.} = 9$	$x_{12.} = 15$	$x_{13.} = 39$	$x_{1..} = 21$
Al- ters- klas- se 2	Versuch 1	$x_{211} = 18$	$x_{221} = 18$	$x_{231} = 6$	
	Versuch 2	$x_{212} = 12$	$x_{222} = 24$	$x_{232} = 12$	
$X_{2j.}$		$X_{21.} = 30$	$X_{22.} = 42$	$X_{23.} = 18$	$X_{2..} = 90$
$x_{2j.}$		$x_{21.} = 15$	$x_{22.} = 21$	$x_{23.} = 9$	$x_{2..} = 15$
$X_{.j.}$		$X_{.1.} = 48$	$X_{.2.} = 72$	$X_{.3.} = 96$	$X_{...} = 216$
$x_{.j.}$		$x_{.1.} = 12$	$x_{.2.} = 18$	$x_{.3.} = 24$	$x_{...} = 18$

Tab. 2 : Ausgangsdaten der zweifachen Varianzanalyse

Es ergeben sich folgende Fragestellungen, die für die Unternehmung von Interesse sind:¹

1. Bestehen Unterschiede zwischen den Absatzchancen der drei Waschmittel? Diese Frage kann durch eine Homogenitätsprüfung der drei Spaltmittelwerte überprüft werden.

¹ Vgl. zu den statistischen Fragestellungen Roßner, R., Zweifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 230ff, hier S. 231.

2. Liegt ein unterschiedliches signifikantes Käuferverhalten bei den drei Waschmitteln in den beiden Hausfrauen-Altersklassen vor? Diese Frage kann durch eine Homogenitätsprüfung der beiden Zeilenmittelwerte überprüft werden.
3. Bestehen Wechselbeziehungen zwischen Waschmittelverwendung und den Altersgruppen der Hausfrauen dergestalt, daß "ältere" Hausfrauen ein bestimmtes Waschmittel intensiver verwenden als "jüngere" Hausfrauen?

Die statistische Fragestellung bei der Zweifachklassifikation besteht darin, analog zur Einfachklassifikation Modellgleichungen abzuleiten, die es gestatten, die Beobachtungswerte x_{ijk} zu erklären.¹ Der Index i ($i = 1, 2, \dots, n$) gibt im obigen Beispiel die jeweilige Altersklasse der Hausfrauen an, der Index j ($j = 1, 2, \dots, m$) identifiziert die verschiedenen Waschmittel und der Index k ($k = 1, 2, \dots, p$) gibt den jeweiligen Versuch bei gegebenem i und j an.

Der Grundgedanke der Zweifachklassifikation besteht darin, die Wirkungen von zwei Faktoren additiv zu erfassen. Es wird dabei zwischen Zeilen- und Spalteneffekten unterschieden. Der Zeileneffekt bezieht sich auf die Differenz zwischen dem jeweiligen Zeilenmittel und dem Gesamtmittel, der Spalteneffekt betrifft die Differenz zwischen dem jeweiligen Spaltenmittel und dem Gesamtmittel. Wechselbeziehungen bestehen dann, wenn die Differenz zwischen zwei Mittelwerten unterschiedlicher Ebene eines Faktors für alle Ebenen des anderen Faktors unterschiedlich ist.

Die obigen drei marketingorientierten Fragestellungen lassen sich in die folgenden statistischen Fragestellungen transformieren:

¹ Die Darstellung erfolgt in Anlehnung an Roßner, R., Zweifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 230-232 und Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 302-307.

1. Liegen bei den Spaltenmitteln Spalteneffekte vor, d.h. sind für die Grundgesamtheit $\beta_j \stackrel{\text{Def.}}{=} \mu_{.j} - \mu_{..}$ anzunehmen?
2. Liegen bei den Zeilenmitteln Zeileneffekte vor, d.h. sind für die Grundgesamtheit $\alpha_i \stackrel{\text{Def.}}{=} \mu_{.i} - \mu_{..}$ anzunehmen?
3. Bestehen Unterschiede zwischen den Wechselwirkungen der Stufe i des Faktors 1 und der Stufe j des Faktors 2, so daß auf Wechselwirkungen in der Grundgesamtheit geschlossen werden darf?

Die Modellgleichung bei der Zweifachklassifikation lautet:¹

$$(13) \quad x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad .$$

In dieser Modellgleichung bedeutet x_{ijk} der i-te Beobachtungswert der j-ten Stichprobe beim k-ten Versuch. Der Zeileneffekt wird durch α_i , der Spalteneffekt durch β_j , die Wechselwirkungen durch $(\alpha\beta)_{ij}$ zum Ausdruck gebracht. Die Größe e_{ijk} bezeichnet den normalverteilten Zufallsfehler beim jeweiligen Beobachtungswert.

Aufgrund der angenommenen Gleichheit der Mittelwerte der Grundgesamtheiten ergibt sich:

$$(14) \quad \sum_i \alpha_i = 0$$

$$(15) \quad \sum_j \beta_j = 0 \quad .$$

Da der Zufallsfehler e_{ijk} normalverteilt ist, gelten die Annahmen

$$(16) \quad \mu(e_{ijk}) = 0$$

$$(17) \quad \text{Var}(e_{ijk}) = \sigma_e^2 \quad .$$

¹ Vgl. Roßner, R., Zweifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 230f.

Darüber hinaus gilt aufgrund von (14) und (15) annahmegeräß

$$(18) \sum_i (\alpha\beta)_{ij} = 0$$

$$(19) \sum_j (\alpha\beta)_{ij} = 0 \quad .$$

Es wird bei der Mehrfachklassifikation analog zur Einfachklassifikation vorgegangen, indem eine Identitätsgleichung für die Abweichung der Beobachtungswerte x_{ijk} vom Gesamtmittel $\bar{x} \dots$ formuliert wird:¹

$$(20) \sum_{ijk} (x_{ijk} - \bar{x} \dots) = \sum_{ijk} (x_{i..} - \bar{x} \dots) + \sum_{ijk} (x_{.j.} - \bar{x} \dots) + \\ \sum_{ijk} (x_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x} \dots) + \\ \sum_{ijk} (x_{ijk} - x_{ij.}) \quad .$$

Durch Quadrierung von (20) und nach einigen Umformungen ergibt sich

$$(21) \sum_{ijk} (x_{ijk} - \bar{x} \dots)^2 = \sum_{ijk} (x_{ijk} - x_{ij.})^2 + p \sum_{ij} (x_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \\ \bar{x} \dots)^2 + p \cdot n \sum_j (x_{.j.} - \bar{x} \dots)^2 + \\ n \cdot m \sum_i (x_{i..} - \bar{x} \dots)^2 \quad .$$

Der Ausdruck auf der linken Seite von (21) gibt die Gesamtsumme der Abweichungsquadrate an. Der erste Term auf der rechten Seite von (21) gibt die Gesamtsumme der Abweichungsquadrate an, die zufallsbedingt sind, der zweite Ausdruck die Gesamtsumme der Abweichungsquadrate, die

¹ Vgl. Ferguson, G.A. Statistical Analysis in Psychology and Education, a.a.O., S. 303.

durch Wechselwirkungen zwischen den beiden Faktoren hervorgerufen werden. Der dritte Ausdruck bestimmt die Gesamtsumme der Abweichungsquadrate, die durch den Faktor j hervorgerufen werden, während schließlich der letzte Ausdruck die Gesamtsumme der Abweichungsquadrate angibt, die durch den Faktor i hervorgerufen werden.

Der nächste, sich bei der Zweifachklassifikation anschließende Schritt, besteht in der Bestimmung der Freiheitsgrade, um die Varianzen berechnen zu können. Die Gesamtzahl der Beobachtungswerte beträgt $N = n \cdot m \cdot p$. Analog zur Einfachklassifikation beträgt der Freiheitsgrad der Gesamtvarianz $N-1$, der Freiheitsgrad der Varianz der Zufallskomponenten $N - m \cdot n$, die Varianz der Wechselwirkungen $(n-1)(m-1)$, die Varianz zwischen den Spalten der Matrix $m-1$ und zwischen den Zeilen $n-1$.

Auch hier gilt, daß die Summe der Abweichungsquadrate bei Mehrfachklassifikation additiv ist, dieses aber nicht bei den Varianzen zutrifft. Die ermittelten Varianzen stellen Schätzwerte für die unbekannten Varianzen der Grundgesamtheit dar. Aufgrund der Zahlenwerte des Beispiels ergibt sich eine Varianz der Zufallskomponenten von 18, eine Varianz der Wechselwirkungen von 432, eine Varianz zwischen den Spalten der Matrix von 144 und eine Varianz zwischen den Zeilen von 108.

Analog zur Einfachklassifikation wird abschließend der F-Test vorgenommen, um die aufgestellten Hypothesen überprüfen zu können. Da der Quotient aus der Varianz zwischen den Spalten und der Varianz der Zufallskomponenten 8 beträgt, der theoretische F-Wert bei 5 Prozent Irrtumswahrscheinlichkeit und den Freiheitsgraden von $f_1 = 2$ und $f_2 = 6$ 5,143 beträgt, ist die Nullhypothese abzulehnen und die Gegenhypothese anzunehmen. Es ergibt sich somit der Schluß, daß Spalteneffekte vorliegen, d.h., daß die verschiedenen Waschmittel unterschiedliche durchschnittliche Absatzchancen haben.

Der Quotient aus der Varianz zwischen den Zeilen der Matrix und der Varianz der Zufallskomponenten beträgt 6. Bei einer Irrtumswahrscheinlichkeit von 5 Prozent und einem Freiheitsgrad von $f_1 = 1$ und $f_2 = 6$ ergibt sich ein theoretischer F-Wert von 5,987. Da der theoretische F-Wert kleiner als der empirische F-Wert ist, muß die Nullhypothese verworfen und die Gegenhypothese angenommen werden. Es bestehen somit Zeileneffekte, d.h. unterschiedliche Altersklassen der Hausfrauen weisen unterschiedliche durchschnittliche Verwendungsintensitäten von Waschmitteln auf.

Schließlich beträgt der Quotient aus der Varianz der Wechselwirkungen und der Varianz der Zufallskomponenten 20,5. Es ergibt sich bei einer Irrtumswahrscheinlichkeit von 5 Prozent und Freiheitsgraden von $f_1 = 2$ und $f_2 = 6$ ein theoretischer F-Wert von 5,143. Da der theoretische F-Wert kleiner als der empirische F-Wert ist, ist die Nullhypothese abzulehnen und die Gegenhypothese anzunehmen. Es bestehen somit Wechselbeziehungen dergestalt, daß jüngere bzw. ältere Hausfrauen jeweils ein spezielles Waschmittel kaufen. Daraus ergibt sich u.a. für das Marketing die Konsequenz, die Waschmittelwerbung zielgruppenorientiert nach der Altersklasse auszurichten.

3.3 Kovarianzanalyse

Aufgabe der Kovarianzanalyse ist die Untersuchung des Zusammenhangs zwischen nominal und metrisch skalierten Faktoren mit dem Ziel, einen Faktor aus der Analyse zu eliminieren.¹ Analog zur Varianzanalyse lassen sich Einfach- und Mehrfachklassifikationen unterscheiden. Die Einfachklassifikation zeichnet sich dadurch aus, daß je ein nominal und metrisch skaliert Faktor vorhanden ist. Bei der Mehrfachklassifikation ist die Zahl der in die Analyse einzubeziehenden nominal und metrisch skalierten Faktoren beliebig. Es muß aber mindestens je ein metrisch und nominal skaliert Faktor vorhanden sein, denn "sonst liegt der Spezialfall eines Modells der Varianzanalyse oder der Regressionsanalyse vor."²

In ökonometrischen Untersuchungen ist es oft notwendig, einen als kausal bedingten Faktor aus der Analyse zu eliminieren. Es muß deshalb eine statistische Methode dafür sorgen, daß die Effekte eines Faktors (Einfachklassifizierung) oder mehrerer Faktoren (Mehrfachklassifizierung) bei Untersuchungen eliminiert werden können, damit eine valide Bewertung der Untersuchungsergebnisse möglich ist.

Ein Beispiel soll die Aufgabenstellung der Kovarianzanalyse bei Einfachklassifikation näher verdeutlichen. Ein Produzent von Futtermitteln hat ein neues Schweinemastfutter entwickelt, von dem erwartet wird, daß es die Gewichtszunahme erheblich beschleunigt. Um die Wirkung des Futtermittels exakt zu überprüfen, werden drei Versuchstiergruppen ($j = 1, 2, 3$) gebildet, von denen zwei mit herkömmlichen Futtermitteln gefüttert werden, während die dritte Versuchstiergruppe mit dem neuen Mastfutter gemästet wird.

1 Vgl. Widdra, W., Einfache Kovarianzanalyse, in Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 324ff, hier S. 325.

2 Derselbe, Mehrfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 329ff, hier S. 329.

Problem ist, die Gewichtszunahme (metrischer Faktor) in Abhängigkeit von den drei Futtermitteln (nominal skalierte Faktoren) zu untersuchen. Das Anfangsgewicht der Versuchstiere vor Beginn des Versuches soll ausgeschaltet werden, um den tatsächlichen Masterfolg feststellen zu können, da die Hypothese plausibel ist, daß die Gewichtszunahme ursprünglich schwerer Versuchstiere größer sein wird als bei ursprünglich leichteren Versuchstieren. Die materielle Fragestellung, die die Kovarianzanalyse lösen soll, besteht darin, ob die Futtermittel hinsichtlich des Masterergebnisses gleichwertig sind oder ob zwischen ihnen signifikante Unterschiede bestehen.

Die Versuchsergebnisse sind in der folgenden Tabelle zusammengestellt, in der x_{ij} das Gewicht eines Versuchstieres zu Beginn des Versuchs und y_{ij} das Gewicht nach Ende des Versuchs angibt.

Die Variable x_{ij} , deren Einfluß es zu eliminieren gilt, wird als Kovariate bezeichnet. Die Kovarianzanalyse bei Einfachklassifikation basiert auf folgenden Annahmen:¹

1. Die Elemente der Grundgesamtheiten, aus denen Stichproben gezogen werden, sind normalverteilt.
2. Es wird Additivität der Effekte eines Faktors vorausgesetzt.
3. Die Stichproben sind als echte Zufallsstichproben voneinander unabhängig, so daß jeder Beobachtungswert als Stichprobe mit normalverteiltem Fehler angesehen werden kann.

¹ Vgl. Widdra, K., Einfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 325.

Grupp- Be- obach- tung i	j = 1		j = 2		j = 3		$X_{.i}$	$Y_{.i}$
	x_1	y_1	x_2	y_2	x_3	y_3		
1	$x_{11} = 3$	$y_{11} = 7$	$x_{12} = 8$	$y_{12} = 15$	$x_{13} = 15$	$y_{13} = 30$	$X_{.1} = 26$	$Y_{.1} = 52$
2	$x_{21} = 5$	$y_{21} = 6$	$x_{22} = 12$	$y_{22} = 24$	$x_{23} = 16$	$y_{23} = 35$	$X_{.2} = 35$	$Y_{.2} = 65$
3	$x_{31} = 7$	$y_{31} = 9$	$x_{32} = 15$	$y_{32} = 25$	$x_{33} = 20$	$y_{33} = 32$	$X_{.3} = 42$	$Y_{.3} = 66$
4	$x_{41} = 9$	$y_{41} = 15$	$x_{42} = 17$	$y_{42} = 26$	$x_{43} = 24$	$y_{43} = 38$	$X_{.4} = 50$	$Y_{.4} = 79$
5	$x_{51} = 11$	$y_{51} = 13$	$x_{52} = 18$	$y_{52} = 30$	$x_{53} = 30$	$y_{53} = 40$	$X_{.5} = 49$	$Y_{.5} = 83$
$X_{.j}$	$X_{.1} = 35$		$X_{.2} = 70$		$X_{.3} = 105$		$X_{..} = 200$	
$Y_{.j}$		$Y_{.1} = 50$		$Y_{.2} = 120$		$Y_{.3} = 175$		$Y_{..} = 345$
$x_{.j}$	$x_{.j} = 7$		$x_{.2} = 14$		$x_{.3} = 21$		$x_{..} = 14$	
$y_{.j}$		$y_{.j} = 10$		$y_{.2} = 24$		$y_{.3} = 35$		$y_{..} = 23$

Tab. 3 : Ausgangsdaten der einfachen Kovarianzanalyse

Die Modellgleichung der Kovarianzanalyse bei Einfachklassifikation lautet:¹

$$(1) \quad y_{ij} \stackrel{\text{Def.}}{=} \mu + \alpha_j + \beta(x_{ij} - \bar{x}_{..}) + e_{ij} \quad .$$

In dieser Gleichung ist y_{ij} der i -te Beobachtungswert der j -ten Stichprobe. Vom Mittelwert $\mu_1 = \mu_2 = \dots = \mu$ wird angenommen, daß sämtliche Grundgesamtheiten über einen gemeinsamen unbekannten Mittelwert verfügen. Als Schätzwert für den unbekannten Mittelwert der Grundgesamtheit wird der Gesamtstichprobenmittelwert genommen. Mit α_j wird der Effekt j in der Grundgesamtheit bezeichnet. Er ist die Differenz zwischen dem jeweiligen Stichprobenmittel und dem unbekannten Mittelwert der Grundgesamtheit, korrigiert um den Einfluß der zu eliminierenden Kovariaten. Der Einfluß der zu eliminierenden Kovariaten wird als Hauptwirkung des Faktors j bezeichnet.

Der Einfluß der zu eliminierenden Kovariaten wird durch den Regressionskoeffizienten β zum Ausdruck gebracht. Er gibt die Steigung der x_{ij}/y_{ij} -Regressionsfunktion an. Mit e_{ij} wird der Zufallsfehler beim i -ten Beobachtungswert in der j -ten Stichprobe bezeichnet.

Analog zur Varianzanalyse gelten auch die Annahmen:

$$(2) \quad \sum_j \alpha_j = 0$$

$$(3) \quad \mu(e_{ij}) = 0$$

$$(4) \quad \text{Var}(e_{ij}) = \sigma_e^2 \quad .$$

¹ Die Darstellung erfolgt in Anlehnung an Widdra, W., Einfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 325f und Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 50.

Die Voraussetzungen für die Anwendbarkeit dieses Modells besteht darin, daß der Wert des Regressionskoeffizienten β von Null verschieden ist und daß sowohl der Regressionskoeffizient als auch die Werte der Kovariaten unabhängig von den Beobachtungswerten des zweiten Faktors sind.¹

Um mittels der Modellgleichung (1) und der in (2) bis (4) enthaltenen Reparameterisierungsbedingungen, die erfüllt sein müssen, um eindeutige Schätzwerte zu erhalten, die Analyse vornehmen zu können, ist es notwendig, von folgenden Identitätsgleichungen auszugehen:

$$(5) \sum_{ij} (x_{ij} - x_{..}) = \sum_{ij} (x_{ij} - x_{.j}) + \sum_{ij} (x_{.j} - x_{..})$$

$$(6) \sum_{ij} (y_{ij} - y_{..}) = \sum_{ij} (y_{ij} - y_{.j}) + \sum_{ij} (y_{.j} - y_{..})$$

Identitätsgleichung (5) besagt, daß die Summe der Abweichungen der x_{ij} vom Gesamtmittelwert gleich ist der Abweichungssumme der x_{ij} vom jeweiligen Stichprobenmittel zuzüglich der Abweichungssumme des jeweiligen Stichprobenmittels vom Gesamtmittelwert. Diese Interpretation gilt analog für Gleichung (6).

Beide Gleichungen können in eine gemeinsame überführt werden, indem sie miteinander multipliziert werden. Es ergibt sich dann

$$(7) \sum_{ij} (x_{ij} - x_{..}) (y_{ij} - y_{..}) = \sum_{ij} (x_{ij} - x_{.j}) (y_{ij} - y_{.j}) + \sum_{ij} (x_{.j} - x_{..}) (y_{ij} - y_{.j}) + \sum_{ij} (x_{ij} - x_{.j}) (y_{.j} - y_{..}) + \sum_{ij} (x_{.j} - x_{..}) (y_{.j} - y_{..})$$

¹ Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 328.

Der zweite Regressionskoeffizient ergibt sich als Zusammenfassung der individuellen Regressionskoeffizienten b_j ($j = 1, 2, \dots, n$) zum Regressionskoeffizienten b_w :¹

$$(10) \quad b_w = \frac{\sum_j \sum_i (x_{ij} - \bar{x}_{.j})(y_{ij} - \bar{y}_{.j})}{\sum_j \sum_i (y_{ij} - \bar{y}_{.j})^2} \quad .$$

Die Zusammenfassung der individuellen Regressionskoeffizienten zum Koeffizienten b_w ist nur unter der Prämisse der Homogenität der individuellen Regressionskoeffizienten möglich.²

Die interessierende statistische Frage besteht nun in der Prüfung, ob die Stichprobenmittel $\bar{x}_{.j}$ statistisch homogen oder signifikant verschieden sind. Sie sind dann statistisch homogen, wenn sie aus einer gemeinsamen Grundgesamtheit stammen, anderenfalls inhomogen.

Bei der Prüfung, ob Unterschiede zwischen den ordinalen Faktoren (Futtermittelarten) bestehen, wird von der Nullhypothese ausgegangen, daß keine Effekte auf der Ebene der Grundgesamtheiten bestehen, so daß alle α_j gleich Null sind. Als Testgröße verwendet man die Beziehung³

$$(11) \quad F^{\text{emp.}} = \frac{\frac{1}{j-1} (R_1 - R_0)}{\frac{1}{j(i-1)-1} R_0} \quad ,$$

1 Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 329.

2 Vgl. ebenda, S. 330.

3 Vgl. Widdra, W., Einfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 326.

Da $\sum_{ij} (x_{ij} - x_{..})(y_{ij} - y_{..}) = m \cdot \sum_{ij} (x_{ij} - x_{..}) \cdot \sum_{ij} (y_{ij} - y_{..})$ ist, und diese Umformung entsprechend für die dritten und vierten Terme auf der rechten Seite von (7) vorgenommen werden kann, ergibt sich nach dieser Umformung und unter Berücksichtigung, daß $\sum_{ij} (x_{ij} - x_{..})$ und $\sum_{ij} (y_{ij} - y_{..})$ gleich Null sind¹

$$(8) \quad \sum_{ij} (x_{ij} - x_{..})(y_{ij} - y_{..}) = \sum_{ij} (x_{ij} - x_{.j})(y_{ij} - y_{.j}) + m \sum_j (x_{.j} - x_{..})(y_{.j} - y_{..})$$

In (8) bedeutet der Term auf der linken Seite die Gesamtproduktsumme der Abweichungen der x_{ij} und der y_{ij} von den jeweiligen Gesamtmitteln. Der erste Term auf der rechten Seite gibt die Abweichungsproduktsumme der x_{ij} bzw. der y_{ij} vom jeweiligen Stichprobenmittel an. Er mißt die Abweichungen innerhalb der jeweiligen Stichprobe, während der dritte Term die Abweichungen zwischen den Stichproben angibt.

Es wird nun die Annahme getroffen, daß der Regressionskoeffizient β bei sämtlichen Stichproben gleich ist.² Es lassen sich zwei verschiedene Regressionskoeffizienten unterscheiden, und zwar der totale Regressionskoeffizient, der es gestattet, die x_{ij} bei Kenntnis der y_{ij} zu prognostizieren:³

$$(9) \quad b_t = \frac{\sum_{ij} (x_{ij} - x_{..})(y_{ij} - y_{..})}{\sum_{ij} (y_{ij} - y_{..})^2}$$

1 Vgl. Widdra, W., Einfache Kovarianzanalyse, in Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 325

2 Vgl. Bock, R.D., Contributions of Multivariate Experimental Designs to Educational Research, in: Cattell, R.B., Hrsg., Handbook of Multivariate Experimental Psychology, Chicago 1966, S. 820ff, hier S. 829.

3 Vgl. Ferguson, G.A., Statistical Analysis in Psychology and Education, a.a.O., S. 329.

in der R_1 die berichtigte Gesamtsumme der Abweichungsquadrate¹

$$(12) R_1 = \sum_{ij} (x_{ij} - x_{..})^2 - \frac{\sum_{ij} (x_{ij} - x_{..})(y_{ij} - y_{..})}{\sum_{ij} (y_{ij} - y_{..})^2}$$

und R_0 die berichtigte Gesamtsumme der Abweichungsquadrate innerhalb der Stichproben ist:²

$$(13) R_0 = \sum_{ij} (x_{ij} - x_{.j})^2 - \frac{\sum_{ij} (x_{ij} - x_{.j})(y_{ij} - y_{.j})}{\sum_{ij} (y_{ij} - y_{.j})^2}$$

Die Testgröße F genügt unter der Bedingung der Nullhypothese einer F -Verteilung mit $j-1$ und $j(i-1)-1$ Freiheitsgraden. Die Nullhypothese ist dann abzulehnen, wenn $F^{\text{emp.}}$ größer ist als $F^{\text{theor.}}_{1-\alpha; j-1, j(i-1)-1}$.

Diese Ausführungen gestatten es, beim einleitend angeführten Beispiel zu entscheiden, ob die Futtermittel hinsichtlich des Masterergebnisses gleichwertig sind (Nullhypothese) oder ob zwischen ihnen signifikante Unterschiede bestehen (Gegenhypothese).

Da der Regressionskoeffizient $b_w = 0,8294$ beträgt, trifft zu, daß er sich von Null unterscheidet. Bei Freiheitsgraden von $j-1=2$ und $j(i-1)-1=11$ ergibt sich bei einer Irrtumswahrscheinlichkeit von 5 Prozent ein $F^{\text{emp.}}_{0,95;2,11}$ vom 12,294. Da der entsprechende theoretische F -Wert 3,982 beträgt, ist die Nullhypothese, daß die Futtermittel hinsichtlich der Masterergebnisse gleichwertig sind, abzulehnen.

¹ Vgl. Widdra, W., Einfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, a.a.O., S. 326.

² Vgl. ebenda.

4. Klassifikations- und Typisierungsverfahren

4.1 Diskriminanzanalyse

Die Diskriminanzanalyse wird dazu verwendet, um eine Trennung zwischen einer a-priori Klassifikation von zwei oder mehreren Gruppen von Merkmalsträgern mit mehreren Variablen auf Basis einer gegebenen Anzahl von Messungen vorzunehmen.¹ Im Gegensatz zu der noch zu erörternden Cluster-Analyse sind die Untersuchungsgruppen bereits gebildet worden. Es geht bei der Diskriminanzanalyse um das Problem, Trennungskriterien zwischen den Gruppen abzuleiten, so daß festgestellt werden kann, wodurch sich ein Element einer bestimmten Gruppe auf Grundlage seiner Variablen von Elementen anderer Gruppen unterscheidet.²

Bei der Diskriminanzanalyse wird angenommen, daß die Grundgesamtheiten der als Stichproben zu interpretierenden Untersuchungsgruppen normalverteilt sind, und daß diesen Grundgesamtheiten zugehörigen Varianzen gleich sind.³

Ein Beispiel soll die Aufgabenstellung der Diskriminanzanalyse näher verdeutlichen. Bei einem neu einzuführenden Produkt werden 12 Testpersonen gebeten, die Produktqualität und die Gebrauchseigenschaften mittels einer zehnstelligen Skala zu bewerten. Die 1 wird einer geringen Produktqualität und schlechten Gebrauchseigenschaften zugewiesen, während die 10 einer hohen Qualität und sehr guten Gebrauchseigenschaften entspricht.⁴ Es werden zwei Gruppen gebildet, wobei in der ersten Gruppe die Testpersonen enthalten sind, die das Produkt als überdurchschnittlich beur-

1 Vgl. Anderson, O., Probleme der statistischen Methodenlehre in den Sozialwissenschaften, 3. Aufl., Würzburg 1957, S. 168.

2 Vgl. Sheth, J.N., Multivariate Analysis in Marketing, a.a.O., S. 36

3 Vgl. dazu Fisher, R.A., The use of multiple measurements in taxonomic problems, in: Annales of Eugenics, Vol. 7, 1937, S. 179ff.

4 Vgl. Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 72.

teilen, in der zweiten Gruppe sich diejenigen Testpersonen befinden, die das Produkt unterdurchschnittlich klassifizieren. Wird aus der Grundgesamtheit aller möglichen Testpersonen eine Versuchsperson herausgegriffen, so stellt sich das Problem, welcher Gruppe sie zugeteilt werden soll und inwieweit die Klassifizierung möglichst weitgehend quantitativ vorgenommen werden kann.¹ Die Untersuchungsergebnisse sind in folgender Tabelle zusammengefaßt:

Test- personen	Qualität x_{1A}	Gebrauchs- eigen- schaften x_{2A}	$x_{1A} - \bar{x}_{1A}$	$x_{2A} - \bar{x}_{2A}$
1	9	8	0,71	1,57
2	7	6	- 1,29	- 0,43
3	10	7	1,71	0,57
4	8	4	- 0,29	- 2,43
5	9	9	0,71	2,57
6	8	6	- 0,29	- 0,43
7	7	5	- 1,29	- 1,43
	$\bar{x}_{1A} = 8,29$	$\bar{x}_{2A} = 6,43$		
	x_{1B}	x_{2B}	$x_{1B} - \bar{x}_{1B}$	$x_{2B} - \bar{x}_{2B}$
8	4	4	0,8	0,2
9	3	6	- 0,2	2,2
10	6	3	2,8	- 0,8
11	2	4	- 1,2	0,2
12	1	2	- 2,2	- 1,8
	$\bar{x}_{1B} = 3,20$	$\bar{x}_{2B} = 3,80$		

Tab. 4 : Ausgangsdaten zur Bestimmung der Diskriminanzfunktion

¹ Vgl. Bräutigam, J. / Küllmer, H., Die Anwendung statistischer Verfahren zur Objektivierung der Kreditwürdigkeitsprüfung, in: Betriebswirtschaftliche Blätter, 21. Jg., 1972, S. 6ff, hier S. 6.

Zur Kennzeichnung der Wertungen der Testpersonen werden die Symbole x_{1A} , x_{2A} , x_{1B} und x_{2B} herangezogen. Diese Merkmale lassen sich auf sehr viele Arten verbinden¹, aus rein pragmatischen Gründen wird eine lineare Trennfunktion gewählt, die im vorliegenden Fall

$$(1) \quad X = b_1 x_1 + b_2 x_2$$

lauten soll. Die Koeffizienten b_1 und b_2 stellen die Gewichte der jeweiligen Merkmalsausprägung dar. Das statistische Problem besteht nun darin, die Koeffizienten b_1 und b_2 so zu bestimmen, daß die Bewertungen in den beiden obigen Testpersonengruppen möglichst genau auseinander gehalten werden können.² Durch Kennzeichnung der ersten Testpersonengruppe mit dem Index A und der zweiten Testpersonengruppe durch den Index B lassen sich die Unterschiede zwischen den Mittelwerten der beiden Testpersonengruppen durch

$$(2) \quad d_1 = \bar{x}_{1A} - \bar{x}_{1B} \quad \text{und} \quad d_2 = \bar{x}_{2A} - \bar{x}_{2B}$$

angeben. Der Gesamtunterschied bestimmt sich durch³

$$(3) \quad d_X = \bar{X}_A - \bar{X}_B = b_1 d_1 + b_2 d_2 \quad .$$

Das Ziel der Diskriminanzanalyse besteht nun darin, die Koeffizienten b der Diskriminanzfunktion (3) derart zu bestimmen, daß der Unterschied zwischen den Mittelwerten der Bewertungsfaktoren möglichst groß und die Summe der Abweichungsquadrate der Bewertungsfaktoren von ihren Mittelwerten möglichst gering ist.⁴ Die zweite Bedingung wird da-

1 Vgl. Bräutigam J. / Küllmer, Die Anwendung statistischer Verfahren zur Objektivierung der Kreditwürdigkeitsprüfung, a.a.O., S. 7.

2 Vgl. Linder, A., Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure, 2. Aufl., Basel-Stuttgart 1957, S. 67.

3 Vgl. ebenda.

4 Vgl. ebenda.

durch zum Ausdruck gebracht, indem

$$(4) \quad T = \sum (x_A - \bar{x}_A)^2 + \sum (x_B - \bar{x}_B)^2$$

ein Minimum wird. Das Wesen der Diskriminanzanalyse liegt nun darin, daß der Ausdruck

$$(5) \quad Z = \frac{d_X^2(b_1; b_2)}{T(b_1; b_2)}$$

maximal wird. In diesem Fall sind die obigen beiden Anforderungen an die Diskriminanzfunktion erfüllt.

Zur Feststellung des Maximums wird (5) partiell nach b_1 und b_2 differenziert. Es ergibt sich bei partieller Ableitung nach b_1

$$(6) \quad \frac{\partial Z}{\partial b_1} = \frac{T \frac{\partial d_X^2}{\partial b_1} - d_X^2 \frac{\partial T}{\partial b_1}}{T^2} = 0 \quad .$$

Dieser Ausdruck kann unter Berücksichtigung, daß

$$\frac{\partial d_X^2}{\partial b_1} = \frac{\partial d_X}{\partial b_1} \cdot 2 d_X \quad \text{ist, in}$$

$$(7) \quad \frac{T}{d_X} \cdot \frac{\partial d_X}{\partial b_1} = \frac{1}{2} \frac{\partial T}{\partial b_1}$$

umgeformt werden. Analog gilt für die partielle Ableitung von (5) nach b_2

$$(8) \quad \frac{T}{d_X} \cdot \frac{\partial d_X}{\partial b_2} = \frac{1}{2} \frac{\partial T}{\partial b_2} \quad .$$

Da $\frac{\partial d_X}{\partial b_1} = d_1$ und $\frac{\partial d_X}{\partial b_2} = d_2$ ist, kann die Diskriminanzfunktion dadurch abgeleitet werden, indem die erste Ableitung von (5) nach b_1 gleich d_1 und die erste partielle Ableitung nach b_2 gleich d_2 gesetzt wird. Der Faktor $\frac{T}{d_X}$ kann dabei vernachlässigt werden, "da die Multiplikation aller b mit einer Konstanten das Verhältnis d_X^2/T nicht verändert...".¹

Werden in (4) die Werte aus der Ausgangstabelle eingesetzt, ergibt sich

$$(9) \quad T = \Sigma |b_1(x_{1A} - \bar{x}_{1A}) + b_2(x_{2A} - \bar{x}_{2A})|^2 + \Sigma |b_1(x_{1B} - \bar{x}_{1B}) + b_2(x_{2B} - \bar{x}_{2B})|^2.$$

Wird (9) partiell nach b_1 und b_2 differenziert, erhält man

$$(10) \quad \frac{\partial T}{\partial b_1} = b_1 \Sigma |(x_{1A} - \bar{x}_{1A})^2 + (x_{1B} - \bar{x}_{1B})^2| + b_2 \Sigma |(x_{2A} - \bar{x}_{2A}) \cdot (x_{1A} - \bar{x}_{1A}) + (x_{2B} - \bar{x}_{2B}) \cdot (x_{1B} - \bar{x}_{1B})| = d_1$$

$$(11) \quad \frac{\partial T}{\partial b_2} = b_1 \Sigma |(x_{1A} - \bar{x}_{1A})(x_{2A} - \bar{x}_{2A}) + (x_{1B} - \bar{x}_{1B})(x_{2B} - \bar{x}_{2B})| + b_2 \Sigma |(x_{2A} - \bar{x}_{2A})^2 + (x_{2B} - \bar{x}_{2B})^2| = d_2$$

Aufgrund der Gleichungen (10) und (11) können die Koeffizienten b_1 und b_2 für die Diskriminanzfunktion bestimmt werden.

Die Diskriminanzfunktion, die aufgrund der Zahlenangaben des Beispiels berechnet wurde, lautet

$$(12) \quad X = 0,2174 x_1 + 0,0308 x_2.$$

¹ Lindner, A., Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure, a.a.O., S. 68.

Aus dieser Funktion läßt sich aufgrund des Gewichts von b_1 feststellen, daß der Produktqualität eine sehr viel höhere Bedeutung zukommt als den Gebrauchseigenschaften.

Die Anwendungsmöglichkeiten der Diskriminanzanalyse sind im Marketing zahlreich. So kann mittels der Diskriminanzanalyse versucht werden, eine Trennung zwischen markentreuen und nichtmarkentreuen Käufern vorzunehmen oder eine Typisierung zwischen verschiedenen Hörer- und Seher-schaften vorzunehmen¹, um nur zwei markante Beispiele aufzuzeigen.²

1 Vgl. dazu Massy, W.F., Discriminant Analysis of Audience Characteristics, in: Journal of Advertising Research, Vol. 5, 1965, S. 39ff.

2 Vgl. ferner die Trennung zwischen Konsumpionieren und übrigen Käufern bei Robertson, S. / Kennedy, J.N., Prediction of Consumer Innovators: An Application of Multiple Discriminant Analysis, in: Aaker, D.A., Hrsg., Multivariate Analysis in Marketing: Theory & Application, Belmont, Calif. 1971, S. 117ff.

4.2 Clusteranalyse

Ziel der Clusteranalyse ist die Ordnung einer Objektmenge entsprechend der ihnen zugehörigen Merkmalsausprägungen in bestimmte sich gegenseitig ausschließende und für sich jeweils umfassende nicht a-priori existierende Objektgruppen.¹ Aus diesem Ziel ergeben sich zwei Unterziele, und zwar erstens die Messung der Ähnlichkeiten zwischen den Objekten und zweitens die Klassifizierung der Objekte entsprechend ihrer Ähnlichkeitsmaße.² Die Zuweisung der Objekte zu bestimmten Objektgruppen soll so erfolgen, daß innerhalb der jeweiligen Gruppen eine möglichst große Ähnlichkeit, zwischen den Gruppen aber ein möglichst großer Unterschied vorhanden ist.³ Die statistische Fragestellung knüpft daran an, welche Messungsprobleme sich bei den Ähnlichkeiten ergeben, und welche Maßgrößen zur Bestimmung der Homogenität der zu klassifizierenden Objekte angewendet werden können. Aus dieser Warte betrachtet, ist die Clusteranalyse ein Gattungsbegriff für statistische Methoden, die die Gruppierung von Objekten betreffen⁴ und keineswegs ein bestimmtes statistisches Verfahren.

Ein Beispiel soll die Aufgabenstellung der Clusteranalyse näher verdeutlichen. Problem sei, eine Gruppe von k Käufern ($k = 1, 2, \dots, s$) in zwei gleichgroße möglichst homogene Gruppen einzuteilen. Die Käufer werden durch jeweils zwei metrisch skalierte Größen, und zwar durch das Alter x_k

1 Vgl. dazu Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 84; ferner Sheth, J.N., The Multivariate Revolution in Marketing Research, a.a.O., S. 16.

2 Vgl. Sheth, J.N., Multivariate Analysis in Marketing, a.a.O., S. 35.

3 Frank, R.E. / Green, P.E., Quantitative Methods in Marketing, a.a.O., S. 85

4 Vgl. Sheth, J.N., Multivariate Analysis in Marketing, a.a.O., S. 35

und das Einkommen y_k charakterisiert. Die Untersuchung hat folgende Ergebnisse erbracht:

Käufer	Merkmalsausprägungen	
	Alter x_k	Einkommen y_k
1	35	2
2	48	7
3	52	6
4	39	4
5	41	3
6	21	1
7	19	5
8	37	8
9	42	7
10	31	10

Tab. 5 : Metrisch skalierte Merkmalsausprägungen der Käufer bei zwei Variablen

Es werden zunächst die absoluten Alters- und Einkommensunterschiede zwischen den Käufern berechnet, die in den beiden folgenden Tabellen angegeben sind:

$x_k \backslash x_k$	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	-									
x_2	13	-								
x_3	17	4	-							
x_4	4	9	13	-						
x_5	6	7	11	2	-					
x_6	14	27	31	18	20	-				
x_7	16	29	33	20	22	2	-			
x_8	2	11	15	2	4	16	18	-		
x_9	7	6	10	3	1	21	23	5	-	
x_{10}	4	17	21	8	10	10	12	6	11	-

Tab. 6 : Absolute Altersunterschiede zwischen den Käufern

y_k	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
y_k										
y_1	-									
y_2	5	-								
y_3	4	1	-							
y_4	2	3	2	-						
y_5	1	4	3	1	-					
y_6	1	6	5	3	2	-				
y_7	3	2	1	1	2	4	-			
y_8	6	1	2	4	5	7	3	-		
y_9	5	0	1	3	4	6	2	1	-	
y_{10}	8	3	4	6	7	9	5	2	3	-

Tab. 7 : Absolute Einkommensunterschiede zwischen den Käufern

Die Clusterbildung erfolgt, indem hier die Euklidische Entfernung zwischen den Kombinationen der Merkmalsausprägungen berechnet wird:

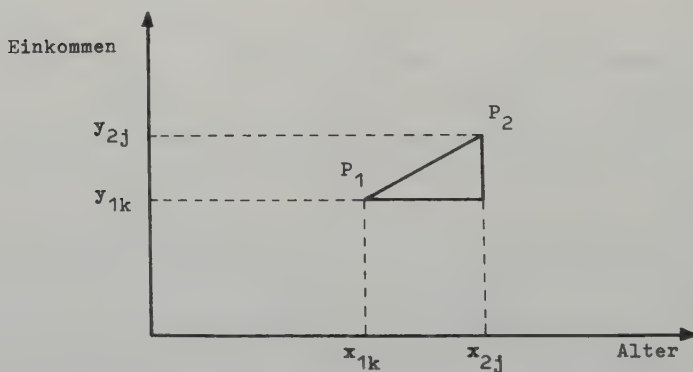


Abb. 12: Berechnung der Entfernung zwischen Kombinationen von zwei Merkmalsausprägungen

Der Abstand zwischen den Punkten P_1 und P_2 errechnet sich durch folgende Formel:

$$(1) \quad E_{ij} = [(x_{2j} - x_{1j})^2 + (y_{2k} - y_{1k})^2]^{\frac{1}{2}}$$

Die Euklidische Entfernung ist der folgenden Tabelle zu entnehmen:

$\begin{matrix} x_k y_k \\ x_k y_k \end{matrix}$	$x_1 y_1$	$x_2 y_2$	$x_3 y_3$	$x_4 y_4$	$x_5 y_5$	$x_6 y_6$	$x_7 y_7$	$x_8 y_8$	$x_9 y_9$	$x_{10} y_{10}$
$x_1 y_1$	-									
$x_2 y_2$	13,93	-								
$x_3 y_3$	17,47	4,10	-							
$x_4 y_4$	4,50	9,50	13,15	-						
$x_5 y_5$	6,10	8,05	11,40	2,30	-					
$x_6 y_6$	14,03	27,66	31,40	18,25	20,10	-				
$x_7 y_7$	16,28	29,07	33,01	20,03	22,10	4,50	-			
$x_8 y_8$	6,30	11,05	15,13	4,50	6,40	17,46	18,24	-		
$x_9 y_9$	8,60	6,00	10,05	4,30	4,10	21,83	23,06	5,10	-	
$x_{10} y_{10}$	10,00	17,26	21,38	10,00	12,20	13,45	13,00	6,30	11,40	-

Tab. 8 : Euklidische Entfernungen bei zwei Merkmalsausprägungen

Als Ausgangspunkt des ersten Clusters wird die Beobachtungswerte-Kombination gewählt, deren Euklidische Entfernung am geringsten ist. Dieses ist bei $x_4 y_4 / x_5 y_5$ der Fall.

Als nächster Schritt schließt sich die Berechnung der Schwerpunkt-Koordinaten der Beobachtungswerte-Kombinationen mit der geringsten Euklidischen Entfernung an. Sie werden mit $\bar{x}_s / \bar{y}_s$ bezeichnet und errechnen sich nach den Formeln $(x_4 + x_5) : 2$ und $(y_4 + y_5) : 2$. Die Beobachtungswerte-Kombinationen werden dann entsprechend der Entfernung zum Schwerpunkt in das erste Cluster aufgenommen, bis die vorgegebene Anzahl

der Beobachtungswert-Kombinationen (hier gleich 5) erreicht ist. Der Abstand der Beobachtungswert-Kombinationen vom Schwerpunkt bestimmt sich aufgrund von

$$(2) \quad [(x_k - \bar{x}_s)^2 + (y_k - \bar{y}_s)^2]^{\frac{1}{2}}.$$

Die Entfernungen zwischen dem Schwerpunkt und den Beobachtungswert-Kombinationen werden durch folgende Tabelle wiedergegeben:

Käufer	$x_k - \bar{x}_s$	$y_k - \bar{y}_s$	$(x_k - \bar{x}_s)^2$	$(y_k - \bar{y}_s)^2$	$(x_k - \bar{x}_s)^2 + (y_k - \bar{y}_s)^2$	$[(x_k - \bar{x}_s)^2 + (y_k - \bar{y}_s)^2]^{\frac{1}{2}}$
1	5	1,5	25	2,25	27,25	5,220
2	8	3,5	64	12,25	76,25	8,732
3	12	2,5	144	6,25	150,25	12,262
4	1	0,5	1	0,25	1,25	1,120
5	1	0,5	1	0,25	1,25	1,120
6	19	2,5	361	6,25	367,25	19,167
7	21	1,5	441	2,25	443,25	21,050
8	3	4,5	9	20,25	29,25	5,405
9	2	3,5	4	12,25	16,25	4,031
10	9	6,5	81	42,25	123,25	11,110

Tab. 9 : Berechnung der Entfernungen zwischen dem Schwerpunkt und Merkmalsausprägungen

Aufgrund der Rechnung ergibt sich, daß die Beobachtungswert-Kombinationen der Käufer 4, 5, 9, 1 und 8 im ersten Cluster, die restlichen Beobachtungswert-Kombinationen im zweiten Cluster zusammengefaßt werden. Die Clusterbildung ist in der folgenden Abbildung verdeutlicht:

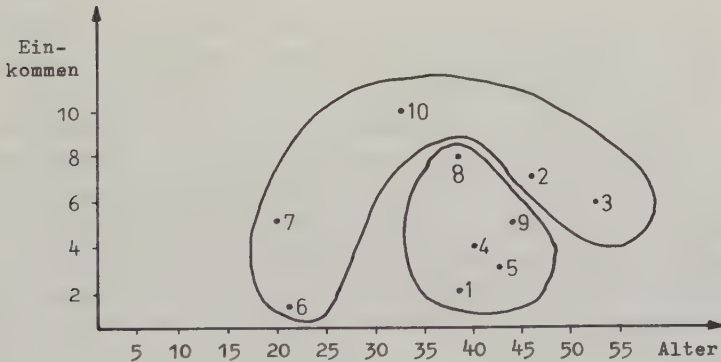


Abb.13 : Clusterbildung bei zwei Merkmalsausprägungen

Die Schwerpunkt-Methode kann dahingehend erweitert werden, daß beliebig viele Merkmalsausprägungen bei der Clusterbildung berücksichtigt werden können. Es ist aber zu beachten, daß bei vielen in die Analyse einbezogenen Merkmalsausprägungen die Objektloadungen interkorreliert sein können, so daß die Merkmalsausprägungen nicht mehr orthogonal sind. In diesem Fall verringert sich die Aussagefähigkeit des Euklidischen Abstandsmaßes erheblich.¹ Ein weiterer Kritikpunkt des erörterten Verfahrens und i.w.S. sämtlicher

¹ Vgl. Frank, R.E. / Green, P.E., Numerical Taxonomy in Marketing Research - A Review Article, in: Journal of Marketing Research, February 1968, S. 83ff, hier S. 86.

Clusterverfahren ist darin zu sehen, daß keine Signifikanztests möglich sind, die darüber Aufschluß geben, ob eine bestimmte Merkmalskombination in ein bestimmtes Cluster gehört oder nicht. Die statistischen Verfahren der Clusterbildung gewährleisten ausschließlich, daß ein einheitlicher Algorithmus zur Clusterbildung zur Anwendung gelangt, während das materielle Problem, ob eine bestimmte Merkmalskombination zu recht in ein berechnetes Cluster gehört, von den zur Verfügung stehenden Algorithmen nicht beantwortet werden kann.

Die Cluster-Bildung aufgrund der Euklidischen Entfernung ist nur eine Möglichkeit von vielen zur Clusterbildung. Es sind weitere Verfahren bekannt.¹ Das einfachste Verfahren, das sich an die Berechnung der Euklidischen Entfernung anschließt, besteht in der Normierung der Merkmalskombinationen mit einem Mittelwert von Null und einer Einheitsvarianz.

Trotz dieser methodischen Nachteile hat sich die Clusteranalyse insbesondere im Zusammenhang mit Problemen der Kunden- und Firmentreue und bei der Klassifikation von Medientypen hinsichtlich ihrer Leser und Seher als nützlich erwiesen.

¹ Vgl. Frank, R.E. / Green, P.E., Numerical Taxonomy in Marketing Research - A Review Article, a.a.O., S. 86ff,

5. Verfahren zur Analyse von Zusammenhängen

5.1 Korrelationsanalyse

Aufgabe der Korrelationsanalyse ist die Untersuchung der Straffheit des Zusammenhangs zwischen den erklärenden Variablen und der erklärten Variable (lineare Mehrfachkorrelation).

Für die Messung der Straffheit des Zusammenhangs sind mehrere Korrelationsmerkmale entwickelt worden. Das einfachste Korrelationsmaß ist der FECHNERSche Korrelationsindex.¹ Dieser Index wird berechnet, indem die Abweichungen der x_i und y_i von ihren Mittelwerten festgestellt werden, wobei es ausschließlich auf die Vorzeichen dieser Abweichungen ankommt. Der FECHNERSche Korrelationsindex mißt den Grad der Abhängigkeit (Korrelation) als Anzahl der Differenz vorzeichenmäßig übereinstimmender und vorzeichenmäßig nicht übereinstimmender Merkmalsausprägungen an der Gesamtzahl der Merkmalsausprägungen. Er liegt zwischen +1 (alle Vorzeichen stimmen überein) und -1 (alle Vorzeichen stimmen nicht überein).

Der Nachteil des FECHNERSchen Korrelationsindex besteht darin, daß er eine Erhöhung bzw. Verminderung der Korrelation nicht erfaßt. Darüber hinaus ist er nur bei linearen Einfachregressionen anwendbar. Als Maß für die Stochastizität ist er nur bedingt brauchbar. Aus diesen Gründen wird ein Korrelationsmaß benötigt, das in Zusammenhang mit den ermittelten Regressionsfunktionen steht und unabhängig von der Anzahl und dem Maßstab der untersuchten Variablen ist. Ein derartiges Maß wird dadurch gewonnen, indem die Abweichungen der Beobachtungswerte von ihren Mittelwerten ($x_i - \bar{x}$) bzw. ($y_i - \bar{y}$) auf ihre jeweiligen Varianzen bezogen werden, so daß die Unabhängigkeit der Maßgröße vom Maßstab und der Anzahl der in die Analyse einbezogenen Variablen gewährleistet ist.

¹ Vgl. Förster, E. / Eggermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 51f.

Es ergibt sich somit¹

$$(1) \quad \sum_i \frac{(x_i - \bar{x})}{s_x} \cdot \frac{(y_i - \bar{y})}{s_y}$$

Beziehung (1) bringt zum Ausdruck, daß die Summe um so größer wird, je mehr die beiden Reihen der Variablen x und y gleich- oder gegenläufig sind.²

Wird (1) durch die Anzahl der Beobachtungswerte dividiert, ergibt sich der BRAVAIS-PEARSONSche Korrelationskoeffizient r_{yx} :³

$$(2) \quad r_{yx} = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})}{s_x \cdot s_y}$$

Da der Zähler von (2) die Kovarianz $s_{y/x}$ ist, kann r_{yx} in der Form

$$(3) \quad r_{yx} = \frac{s_{x/y}}{s_x \cdot s_y}$$

geschrieben werden. Der BRAVAIS-PEARSONSche Korrelationskoeffizient ist somit das Verhältnis zwischen der Kovarianz zum Produkt der Standardabweichungen der Variablen x und y. Werden für diese Standardabweichungen die entsprechenden Formeln in (2) eingesetzt, ergibt sich für den Korrelationskoeffizienten:

$$(4) \quad r_{yx} = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i [(x_i - \bar{x})^2 (y_i - \bar{y})^2]}}$$

¹ Vgl. Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 116.

² Vgl. ebenda.

³ Vgl. ebenda, S. 116ff.

Dieser Korrelationskoeffizient liegt zwischen +1 und -1, wobei er bei einer positiven Korrelation positiv und bei einer negativen Korrelation negativ ist. Bei Multiplikation von r_{yx} mit 100 gibt der so erhaltene Ausdruck den Prozentsatz der Streuung an, der durch die Regressionsfunktion erklärt wird. Der Rest ist nicht erklärte Zufallsstreuung.

Darüber hinaus ermöglicht es die Korrelationsanalyse, die Straffheit des Zusammenhangs zwischen mehr als zwei Variablen zu untersuchen. Für diesbezügliche Analysen bietet sich der mehrfache oder multiple Korrelationskoeffizient an. Bei dreifach linearer Korrelation mit der zu erklärenden Variable y und den zu erklärenden Variablen x und z gibt der multiple Korrelationskoeffizient die Straffheit des Zusammenhangs unter der Bedingung an, daß die Variable y gleichzeitig von den erklärenden Variablen x und z abhängt.

Der multiple Korrelationskoeffizient wird mit $r_{y/xz}$ bezeichnet. Er lautet in allgemeiner Schreibweise:

$$(5) \quad r_{y/xz}^2 = \frac{\sum_1 (y'_i - \bar{y})^2}{\sum_1 (y_i - \bar{y})^2}$$

Wird die bei der Regressionsanalyse abgeleitete Beziehung

(8) $y'_i - \bar{y} = b(x_i - \bar{x}) + c(z_i - \bar{z})$ in (5) eingesetzt und beachtet, daß $\sum_1 (y_i - \bar{y})^2$ gleich s_y^2 ist, ergibt sich

$$(6) \quad r_{y/xz}^2 = \frac{\sum_1 [b(x_i - \bar{x}) + c(z_i - \bar{z})]^2}{s_y^2}.$$

Wird die Quadratur des Zählers vorgenommen, erhält man folgenden Ausdruck

$$(7) \quad r_{y/xz}^2 = \frac{\sum_1 [b^2(x_i - \bar{x})^2 + 2bc(x_i - \bar{x})(z_i - \bar{z}) + c^2(z_i - \bar{z})^2]}{s_y^2}.$$

der in

$$(8) \quad r_{y/xz}^2 = \frac{\bar{y} [b^2 \cdot s_x^2 + 2bc \cdot s_{x/z} + c^2 \cdot s_z^2]}{\bar{y}^2}$$

umgeformt werden kann.

Ausdruck (8) kann durch weitere Umformungen, auf die hier verzichtet wird, in¹

$$(9) \quad r_{y/xz}^2 = \frac{r_{yx}^2 + r_{yz}^2 - 2r_{yx} \cdot r_{yz} \cdot r_{xz}}{1 - r_{xz}^2}$$

umgeformt werden. Es gilt sowohl beim einfachen als auch beim multiplen Korrelationskoeffizienten, daß $r_{yx} = r_{xy}$ und daß $r_{y/xz} = r_{x/yz} = r_{z/yx}$ ist. Es genügt somit, einen Korrelationskoeffizienten zu berechnen. Bei multiplen Korrelationskoeffizienten wird man zweckmäßigerweise so vorgehen und zunächst die einfachen Korrelationskoeffizienten ermitteln, um dann die entsprechende Berechnung des multiplen Korrelationskoeffizienten vorzunehmen.

¹ Vgl. dazu die Ableitung bei Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 127f.

5.2 Faktorenanalyse

"Das Hauptziel der Faktorenanalyse ist die Ableitung hypothetischer Größen oder Faktoren aus einer Menge beobachteter Variablen. Die Faktoren sollen möglichst einfach sein und die Beobachtungen hinreichend genau beschreiben und erklären."¹

Als Faktoren werden dabei mathematische Größen verstanden, die aus den beobachteten nominal, ordinal oder metrisch skalierten Beobachtungsdaten extrahiert und interpretiert werden. Aus den Aufgabenstellungen, denen die Faktorenanalyse genügt, lassen sich folgende Primäraufgaben herausgreifen:²

1. Die originäre Aufgabe der Faktorenanalyse besteht darin, aus einem großen Beobachtungsdatenvolumen Faktoren zu extrahieren, die geeignet sind, ohne allzu großen Informationsverlust den Informationsgehalt der Beobachtungsdaten durch wenige Faktoren wiederzugeben. Aufgrund dieser originären Aufgabenstellung ist die Faktorenanalyse in der Lage, die Basisstruktur von Beobachtungsdaten offenzulegen, um bislang unbekannte Phänomene anhand weniger Einflußgrößen untersuchen zu können, um eine empirische Typologie zu entwerfen.
2. Eine weitere wesentliche Aufgabe der Faktorenanalyse besteht aufgrund der mit ihr verbundenen Reduzierung von Beobachtungsdaten darin, daß Hypothesen über komplexe reale Phänomene aufgestellt und überprüft werden können.

1 Überla, K., Faktorenanalyse, a.a.O., S. 3. (Im Original kursiv.)

2 Vgl. dazu Rummel, R.J., Understanding factor analysis, in: The Journal of Conflict Resolution - A Quarterly for Research Related to War and Peace, Vol. XI, No. 4, 1967, S. 444ff, hier S. 448ff.

3. Weiterhin können die aus einer Datenfülle extrahierten Faktoren sehr viel besser interpretiert werden als eine Vielzahl miteinander korrelierter Daten.¹

Diese drei Ebenen der Faktorenanalyse, die algebraisch-rechnerische Seite, die statistische Seite und die Interpretation sind im folgenden zu behandeln.²

Die Faktorenanalyse ist kein eindeutiges mathematisch-statistisches Verfahren, sondern sie umfaßt eine Vielzahl von Techniken, die oft nur Näherungslösungen darstellen. Aus diesem Grund ist mit der praktischen Vornahme der Faktorenanalyse ein extrem hoher Rechenaufwand verbunden, daß bei vielen in die Analyse eingehenden Variablen, Lösungen nur mittels der elektronischen Datenverarbeitung gefunden werden können.³

Die Faktorenanalyse beginnt mit der Erstellung einer Datenmatrix⁴, in der die bei j Merkmalsträgern ($j=1,2,\dots,n$) festgestellten i Variablen ($i=1,2,\dots,m$) erfaßt werden. Diese $m \times n$ Matrix wird zeilenweise nach der Beziehung $(y_{ij} - \bar{y})/s_i$ standardisiert, so daß sich eine standardisierte Matrix der Ausgangsdaten ergibt, deren Zeilen einen Mittelwert von Null und die Varianz von 1 besitzen.

Der nächste sich bei Vornahme der Faktorenanalyse anschließende Schritt besteht in der Berechnung der Korrelationskoeffizienten zwischen jeweils zwei Variablen. Die Korrelationskoeffizienten werden in der Korrelationskoeffizientenmatrix R mit den Elementen r_{ij} zusammengefaßt. Die Be-

1 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 5.

2 Vgl. ebenda.

3 Vgl. ebenda, S. 4.

4 Vgl. zu den folgenden Ausführungen Überla, K., Faktorenanalyse, a.a.O., S. 62 und insbesondere S. 50ff.

rechnung der Korrelationskoeffizienten aufgrund der standardisierten Datenmatrix hat den Vorteil, daß sich erhebliche rechentechnische Vereinfachungen ergeben.

Der Grundgedanke der Faktorenanalyse ist zunächst rein mathematischer Natur. Sie setzt mit der mathematischen Überlegung ein, wie sich die Matrix der Korrelationskoeffizienten durch eine zweite Matrix möglichst weniger aber hoch geladener Faktoren abbilden bzw. reduzieren läßt. Anders formuliert lautet die Grundfragestellung der Faktorenanalyse: Wie kann eine Reproduktion der Korrelationskoeffizientenmatrix durch eine unbekannte zweite Matrix weniger aber hoch geladener Faktoren und unter Beachtung der Tatsache vorgenommen werden, daß in der Hauptdiagonalen der Korrelationskoeffizientenmatrix aufgrund der Rechenregeln der Matrizenmultiplikation zu schätzende Werte stehen. Dieses für das Verständnis der Faktorenanalyse notwendige Grundtheorem soll im folgenden näher verdeutlicht werden.

Da durch Standardisierung der Datenmatrix, die mit Y_{ij} bezeichnet wird,

$$(1) \quad \begin{pmatrix} y_{11} & \dots & y_{1j} & \dots & y_{1n} \\ \vdots & & \vdots & & \vdots \\ y_{i1} & \dots & y_{ij} & \dots & y_{in} \\ \vdots & & \vdots & & \vdots \\ y_{m1} & & y_{mj} & & y_{mn} \end{pmatrix} = Y_{ij}$$

sich folgende Matrix ergibt,

$$(2) \quad \begin{pmatrix} \frac{y_{11} - \bar{y}_1}{s_1} = z_{11} \dots \frac{y_{1j} - \bar{y}_j}{s_j} = z_{1j} \dots \frac{y_{1n} - \bar{y}_n}{s_n} = z_{1n} \\ \vdots \\ \frac{y_{i1} - \bar{y}_i}{s_i} = z_{i1} \dots \frac{y_{ij} - \bar{y}_j}{s_j} = z_{ij} \dots \frac{y_{in} - \bar{y}_n}{s_n} = z_{in} \\ \vdots \\ \frac{y_{m1} - \bar{y}_m}{s_m} = z_{m1} \dots \frac{y_{mj} - \bar{y}_m}{s_m} = z_{mj} \dots \frac{y_{mn} - \bar{y}_m}{s_m} = z_{mn} \end{pmatrix} = Z_{ij}$$

bei der $\frac{1}{n} \sum_j z_{ij}^2 = 1$ ist, kann die Formel für den Korrelationskoeffizienten

$$(3) \quad r_{ik} = \frac{\sum_j (y_{ij} - \bar{y}_i)(y_{kj} - \bar{y}_k)}{\sqrt{\sum_j (y_{ij} - \bar{y}_i)^2 \sum_j (y_{kj} - \bar{y}_k)^2}} \quad \text{bei } (i, k=1, 2, \dots, m)$$

in

$$(4) \quad r_{ik} = \sum_j z_{ij} z_{kj}$$

umgeformt werden. In Matrizenschreibweise lautet dann die Matrix R der Korrelationskoeffizienten

$$(5) \quad R_{ij} = Z_{ij} \cdot Z'_{ij} \quad ,$$

wobei die Matrix Z'_{ij} die transponierte Matrix von Z_{ij} darstellt.¹ "Es ist das Ziel jeder Faktorenanalyse, den Wert z_{ij} , also ein Element von Z, als eine lineare Kombination mehrerer hypothetischer Variablen oder Faktoren zu beschreiben."² Unter der Annahme, daß z_{ij} als Linearkombina-

¹ Vgl. hierzu Überla, K., Faktorenanalyse, a.a.O., S. 24f.

² Ebenda, S. 51.

tion von r Faktoren ausgedrückt werden soll, ergibt sich

$$(6) \quad z_{ij} = a_{i1}F_{1j} + a_{i2}F_{2j} + \dots + a_{ir}F_{rj}.$$

In dieser Funktion sind die a_{i1} bis a_{ir} Koeffizienten, die es zu ermitteln gilt. Die z_{ij} sind die Korrelationskoeffizienten aus der Matrix (5), und die F_{1j} bis F_{rj} stehen für bestimmte Merkmalsträger ($j=1,2,\dots,n$) mit den Faktorenwerten F_{1j} bis F_{rj} . "It is crucial in understanding factor analysis to remember that F stands for a function of variables and not a variable."¹

In Matrizenform läßt sich für alle z_{ij}

$$(7) \quad \begin{pmatrix} z_{11} & \dots & z_{1j} & \dots & z_{1n} \\ \vdots & & \vdots & & \vdots \\ z_{i1} & \dots & z_{ij} & \dots & z_{in} \\ \vdots & & \vdots & & \vdots \\ z_{m1} & \dots & z_{mj} & \dots & z_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{il} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} F_{11} & \dots & F_{1j} & \dots & F_{1n} \\ \vdots & & \vdots & & \vdots \\ F_{i1} & \dots & F_{ij} & \dots & F_{in} \\ \vdots & & \vdots & & \vdots \\ F_{r1} & \dots & F_{rj} & \dots & F_{rn} \end{pmatrix}$$

schreiben, wobei $l=1,2,\dots,r$ den Faktorindex angibt. Es gilt somit, daß

$$(8) \quad Z_{ij} = A_{il} \cdot F_{lj}$$

ist. Die Matrix A gilt es zu ermitteln. Sie wird als Faktormuster bezeichnet, wobei ihre Spalten die gesuchten Faktoren darstellen und die jeweiligen a_{il} -Werte darüber Auskunft geben, mit welchem Gewicht (Faktorladung) der Faktor l am Auftreten der beobachteten Variablen beteiligt ist. "Die Matrix Z wird... gewissermaßen in zwei Teile... A und F zerlegt. Dabei enthält A die Beziehungen der Variablen zu den Faktoren und .. F beschreibt die einzelnen.. [Untersuchungsobjekte]."² Problem ist, daß die Matrizen A

¹ Rummel, R.J., Understanding factor analysis, a.a.O., S. 459. (Im Original kursiv)

² Überla, K., Faktorenanalyse, a.a.O., S. 52.

und F unbekannt sind.

Wird $Z_{ij} = A_{i1} \cdot F_{1j}$ in (5) eingesetzt, ergibt sich

$$(9) \quad R_{ij} = A_{i1} \cdot F_{1j} \cdot Z'_{ij} \\ = A_{i1} \cdot F_{1j} (A_{i1} \cdot F_{1j})' .$$

Nach den Regeln der Matrizenmultiplikation¹ und unter Beachtung, daß die zu extrahierenden Faktoren unkorreliert sein sollen, erhält man aus (9)

$$(10) \quad R_{ij} = A_{i1} \cdot A'_{i1} .$$

Dieser Ausdruck wird als Fundamentaltheorem der Faktorenanalyse bezeichnet. Er besagt, daß die Korrelationsmatrix R_{ij} sich durch das Faktorenmuster A_{i1} und die Korrelationsmatrix zwischen den Faktoren A'_{i1} reproduzieren läßt.

Problem ist nun die Bestimmung des unbekannten Faktoren-musters A_{i1} dergestalt, daß sich die Korrelationsmatrix R_{ij} durch das Faktorenmuster A_{i1} und durch die Korrelationsmatrix zwischen den Faktoren A'_{i1} reproduzieren läßt. "Die Faktorenanalyse stellt nun die Frage, ob es eine Größe gibt, einen sog. Faktor, der in Verbindung mit einer geeigneten Rechenregel die beobachteten Korrelationen möglichst genau zu reproduzieren gestattet. Dieser Faktor und die damit verbundene Rechenregel ist zunächst hypothetisch."²

Aufgrund mathematischer Regeln der Matrizenmultiplikation folgt, daß in der Hauptdiagonalen der reduzierten Korrelationsmatrix andere Werte stehen als in der Korrelationsmatrix zwischen den Faktoren. In der Hauptdiagonalen der reduzierten Korrelationsmatrix stehen die Korrelationen von Variablen mit sich selbst, so daß die Hauptdiagonale

¹ Vgl. dazu Überla, K., Faktorenanalyse, a.a.O., S. 22ff.
² Ebenda, S. 44.

mit jeweils der eins besetzt ist. Bei Multiplikation des Faktormusters mit der Korrelationsmatrix zwischen den Faktoren ergibt sich, daß die Werte in der Hauptdiagonalen des Faktormusters nicht mehr mit den Werten in der Hauptdiagonalen der reduzierten Korrelationsmatrix übereinstimmen.

Das erste Problem, das sich bei Vornahme einer Faktorenanalyse stellt, besteht darin, die Werte in der Hauptdiagonalen des Faktormusters zu schätzen. Diese Schätzwerte werden als Kommunalitäten bezeichnet. Die Lösung des Kommunalitätenproblems hat vor der Extraktion von Faktoren zu erfolgen.¹ Erst, wenn die Werte der Kommunalitäten festliegen, sind sämtliche Elemente der reduzierten Korrelationsmatrix bestimmt, so daß damit auch der Rang dieser Matrix festliegt, "d.h. die minimal notwendige Anzahl der Faktoren im mathematischen Sinne."²

Für die Lösung des Kommunalitätenproblems existieren zahlreiche Lösungsverfahren.³ Die drei bedeutendsten Verfahren sind die Methode des höchsten Korrelationskoeffizienten, die Methode des Quadrats des multiplen Korrelationskoeffizienten und die Iteration.

Die einfachste, wenn auch logisch mit den "wahren" Kommunalitäten nicht zusammenhängende Lösungsmöglichkeit, bietet die Methode des höchsten Korrelationskoeffizienten. Diese Methode ist so angelegt, daß aus jeder Spalte der reduzierten Korrelationsmatrix das Element mit dem höchsten Wert bestimmt und in die Hauptdiagonale eingesetzt wird. Obwohl dieses Verfahren einfach ist und brauchbare Ergebnisse erbracht hat, ist es willkürlich und theoretisch nicht abzusichern.

1 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 155.

2 Ebenda, S. 156f.

3 Vgl. dazu ebenda, S. 157 und die dort angeführte Literatur.

Die zweite Methode zur Berechnung der Kommunalitäten - das Quadrat des multiplen Korrelationskoeffizienten - hat materiellen Bezug zur Kommunalität. Das Quadrat des multiplen Korrelationskoeffizienten gibt den Teil der Varianz einer Variable an, den diese mit den restlichen Variablen gemein hat.¹ Wird die Kommunalität nach dieser Methode berechnet, so ist die Differenz zwischen der Kommunalität (die Varianz der Variable i , die sie mit den Faktoren gemein hat) und der Varianz der Variable i diejenige gemeinsame Varianz, die der Variable i zugeordnet wird.²

Die Berechnung der Matrix der multiplen Korrelationskoeffizienten erfolgt am günstigsten mittels Gaußscher Multiplikatoren.³ Dabei wird zunächst die Matrix der reduzierten Korrelationskoeffizienten in R^{-1} invertiert. Da die Beziehung $R \cdot R^{-1}$ gilt, läßt sich die inverse Matrix der Korrelationskoeffizienten mittels der Methode der Gaußschen Multiplikatoren berechnen. Die multiplen Korrelationskoeffizienten ergeben sich dann als die Diagonalelemente der invertierten Korrelationsmatrix. Die Bestimmung der Kommunalitäten erfolgt anschließend, indem die Quadratsumme der Faktorladungen über den jeweiligen Variablen gebildet wird.

Das bekannteste Verfahren zur Bestimmung von Kommunalitäten ist das der Iteration.⁴ Die Methode ist so konzipiert, daß eine a-priori Schätzung der Anzahl der zu extrahierenden Faktoren vorgenommen wird, und dann die Kommunalitäten nach der Methode des Quadrats des multiplen Korrelationskoeffizienten berechnet werden. Mit diesen in die Haupt-

1 Vgl. Rummel, R.J., Understanding factor analysis, a.a.O., S. 461f.

2 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 158.

3 Zur Methode der Gaußschen Multiplikatoren vgl. z.B. Förster, E. / Eggermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 94ff.

4 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 159ff.

diagonale eingesetzten Kommunalitäten wird dann eine festgelegte Anzahl von Faktoren extrahiert, um eine erneute Berechnung der Kommunalitäten vornehmen zu können, um sie dann erneut in die Hauptdiagonale einzusetzen. Dieser Prozeß vollzieht sich so lange, bis die Differenz zwischen eingesetzten und berechneten Kommunalitäten minimal erscheint.

Der nächste bei der Faktorenanalyse zu vollziehende Schritt besteht in der Extraktion der Faktoren. Hierbei geht es um das Problem, wie viele Faktoren aus den vorhandenen Variablen extrahiert werden sollen. Es existieren für die Faktorextraktion mehrere Verfahren, von denen die Hauptkomponentenmethode und die Zentroidmethode die gängigsten Extraktionsverfahren sind.

Der Grundgedanke der Hauptkomponentenmethode besteht darin, die längste Achse zu bestimmen, die durch den Schwerpunkt der Beobachtungspunktwolke führt. Im dreidimensionalen Fall verläuft die erste Hauptachse parallel zur y/x -Ebene, die zweite Hauptachse parallel zur y/z -Ebene.¹ "Durch die Hauptachsenmethode werden die Richtungen der Hauptachsen in bezug auf das ursprüngliche Koordinatensystem festgelegt."² Bei Annahme normalverteilter Zufallsvariablen entstehen bei Schnitten parallel zur x/y -Ebene und parallel zur y/z -Ebene Ellipsen.³ Die Hauptachsen dieser Ellipsen entsprechen dann den zu extrahierenden Faktoren. Die Kommunalitäten (h) werden als Quadrat des multiplen Korrelationskoeffizienten berechnet und sind definiert als

$$(11) \quad h_j^2 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 .$$

1 Vgl. hierzu die Abbildungen bei Überla, K., Faktorenanalyse, a.a.O., S. 94f.

2 Ebenda, S. 95.

3 Vgl. Förster, E. / Eggermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 222.

Das Ziel der Hauptkomponentenmethode besteht nun darin, zunächst den Faktor F_1 so zu extrahieren, daß sein Beitrag zur Kommunalität h_j^2 jeder Variable ein Maximum wird. Ausgehend von der Grundgleichung der Faktorenanalyse

$$(12) \quad \begin{array}{l} z_{11} = a_{11} F_{11} + a_{12} F_{12} + \dots + a_{1n} F_{1n} \\ \vdots \\ z_{i1} = a_{i1} F_{11} + a_{i2} F_{12} + \dots + a_{in} F_{1n} \\ \vdots \\ z_{m1} = a_{m1} F_{11} + a_{m2} F_{12} + \dots + a_{mn} F_{1n} \end{array}$$

ist dafür zu sorgen, daß $\sum_{i=1}^m a_{i1}^2$ ein Maximum wird. Ist der erste Faktor extrahiert worden, so ist beim zweiten Faktor sicherzustellen, daß $\sum_{i=1}^m a_{i2}^2$ ein Maximum wird. Es werden nun so lange Faktoren extrahiert, bis ein vorgegebener Prozentsatz der Kommunalität durch sie erklärt worden ist. Diese Vorgehensweise gewährleistet, daß die extrahierten Faktoren stets ein Minimum der noch verbleibenden Varianz erklären. Dieses Verfahren wird nun so lange fortgesetzt, bis die Residualmatrix im Idealfall nur noch aus Nullladungen besteht. Dabei kann es allerdings geschehen, daß die Anzahl der extrahierten Faktoren gleich der Anzahl der in die Analyse einbezogenen Variablen ist. Aus diesem Grund wird die Faktorextraktion aus Praktikabilitätsgründen dann abgebrochen, wenn durch die Anzahl der extrahierten Faktoren ein vorgegebener Prozentsatz der Kommunalität erklärt wird.¹

Das Grundprinzip der Hauptkomponentenmethode basiert auf der Vorstellung, "daß alle standardisierten Variablen z_j ($j=1,2,\dots,n$) einen n -dimensionalen Raum aufspannen und im Fall der Normalverteilung dieser Variablen die entsprechend einfachen Korrelationskoeffizienten auf einem n -dimensionalen Ellipsoid liegen."² Diese Ellipsoide ergeben sich aufgrund von Schnitten parallel zur x/y -Achse

¹ Vgl. Jahn, W. / Vahle, H., Die Faktoranalyse, a.a.O., S. 107f.

² Ebenda, S. 92.

durch die Dichtefunktion einer stetig zufällig normalverteilten Veränderlichen. Sie liegen konzentrisch zueinander, und ihre Hauptachsen verlaufen parallel zu den Koordinaten der Schnittebene. Die x/y -Regression ist dabei die erste Hauptachse des Ellipsoids, während die y/x -Regression die zweite Hauptachse bildet. Beide Regressionsfunktionen schneiden sich im Schwerpunkt, der der Mittelpunkt sämtlicher konzentrisch zueinander liegenden Ellipsoide ist.¹ Der Rechenaufwand, der bei der Extraktion von Faktoren bei mehreren Variablen anfällt, ist beträchtlich, so daß es bei umfangreichen Extraktionsproblemen notwendig ist, sich Computerprogramme zu bedienen, die es bereits heute ermöglichen, faktoranalytische Extraktionen der Hauptachsen bei 150 Variablen und mehr vorzunehmen.²

Das nächste gängige Verfahren zur Faktorextraktion ist die Zentroidmethode. Sie baut auf der Überlegung auf, daß eine Koordinatenachse festgelegt werden soll, die durch den Schwerpunkt der Punktwolke führt. Die zweite Hauptachse wird durch das Postulat, daß sie zur ersten orthogonal verlaufen soll, bestimmt.

Während bei der Hauptkomponentenmethode eine Lösung zu bestimmen ist, bei der ein Koordinatensystem berechnet werden muß, entlang dessen erster Hauptachse ein Minimum der Varianz der Variablen liegen sollte, wird bei der Zentroidmethode die Drehung des Ausgangskordinatensystems gefordert, so daß eine Koordinatenachse durch den Schwerpunkt der Punktwolke führt. Dieser Schwerpunkt zeichnet sich dadurch aus, daß die Abweichungssumme der Beobachtungswerte von den Schwerpunktkoordinaten gleich Null ist.³

1 Vgl. dazu Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 222ff.

2 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 107.

3 Vgl. Jahn, W. / Vahle, H., Die Faktorenanalyse, a.a.O., S. 72ff.

Hauptfaktoren- und Zentroidmethode liefern, je nachdem welche Rechenverfahren im einzelnen zur Anwendung gelangen, verschiedene Faktorstrukturen. Darüber hinaus sind die erhaltenen Lösungen oft nur schwer zu interpretieren. "Das liegt daran, daß die erhaltene Faktorstruktur entsprechend der Konstruktion der Lösungsverfahren im ersten Faktor alle hohe Ladungen enthält."¹ Durch die Faktorrotation soll erreicht werden, daß die extrahierten Faktoren interpretierbar werden und vermieden werden, daß man eine Faktorstruktur bekommt, die einen Faktor enthält, dessen Ladung in allen Variablen stark vorhanden ist.² Es läßt sich durch mathematische Überlegungen zeigen, daß diese Faktorstrukturen keine eindeutige Lösung darstellen. Sie sind "bis auf eine 'Rotation', das heißt eine orthogonale Transformation oder ganz allgemein eine affine Transformation der in einem kartesischen Koordinatensystem dargestellten Variablen bestimmt."³ Bei der Faktorrotation lassen sich orthogonale und schiefwinklige Rotationen unterscheiden.

Bei der orthogonalen Faktorrotation geht es darum das Ausgangskordinatensystem so zu drehen, daß sich ein Faktormuster ergibt, das möglichst viele hoch und möglichst wenige gering geladene Faktoren aufweist. Bei dieser Rotation handelt es sich geometrisch um eine affine Abbildung der Ausgangskordinaten auf ein neues, um den Ursprung des Ausgangskordinatensystems gedrehtes Koordinatensystem. Durch Bestimmung des Rotationswinkels lassen sich die Koordinaten des rotierten Koordinatensystems berechnen. Diese orthogonale Rotation ist in folgender Abbildung dargestellt:⁴

1 Vgl. Jahn, W. / Vahle, H., Die Faktoranalyse, a.a.O., S. 111.

2 Vgl. ebenda,

3 Ebenda, S. 113.

4 Quelle: Überla, K., Faktorenanalyse, a.a.O., S. 170.

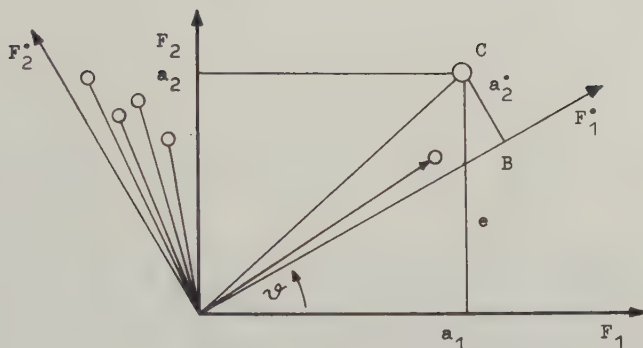


Abb. 14: Orthogonale Faktorrotation

Diese Abbildung soll verdeutlichen, daß es bei der orthogonalen Rotation darum geht, die Punkte a_1 und a_2 auf das neue Koordinatensystem abzubilden. Es wird angenommen, daß der Rotationswinkel v bekannt ist. Durch die beiden Abbildungsgleichungen

$$(13) \quad a_1^* = a_1 \cos v + a_2 \sin v$$

$$(14) \quad a_2^* = a_1 \sin v + a_2 \cos v$$

lassen sich die bekannten Punkte a_1 und a_2 auf das neue Koordinatensystem abbilden.

Bislang wurde der Rotationswinkel v als bekannt vorausgesetzt. Es erhebt sich die Frage, durch welche Überlegungen dieser Winkel bestimmt werden kann. Bei einer Rotation der Faktoren F_1 und F_2 im Ursprung um den Winkel v bleiben die Entfernungen zwischen den Kombinationen von Faktorladungen (hier a_1 und a_2) konstant, ebenfalls die Winkel zwischen den als Vektoren dargestellten Faktorladungskombinationen. Daraus folgt, daß sich eine gegebene Faktorstruktur durch

beliebig viele ineinander überführbare Koordinatenlagen mathematisch äquivalent beschreiben läßt, da die Merkmalskonfigurationen unbeschadet der vorgenommenen Abbildung gleich bleiben.¹ "Dies ist der Grund für die Vieldeutigkeit der Faktorstrukturen."²

Um sich dem Dilemma zu entziehen, daß die Faktorstruktur bis auf unendlich viele affine Abbildungen bestimmt ist, wurden von Thurstone Prinzipien der Einfachstruktur entwickelt, die dafür sorgen sollen, daß das rotierte Koordinatensystem die Variablen möglichst einfach beschreibt.³ Die Prinzipien der Einfachstruktur umfassen dabei folgende Gesichtspunkte:⁴

1. Jede Zeile der Faktormatrix soll wenigstens eine Nullladung aufweisen. Daraus folgt, daß jede Variable mindestens in einer Koordinatenebene liegt.
2. Bei m gemeinsamen Faktoren soll jede Spalte wenigstens m Nullen aufweisen. Daraus ergibt sich, daß jeder Faktor in mindestens m Merkmalen keinen gesicherten Einfluß ausübt, so daß er durch m Variablen in seiner Koordinatenebene lokalisiert werden kann.
3. Für jedes Spaltenpaar sollen die Ladungen einer Zeile in einer Spalte eine Nullladung, in der anderen Spalte eine substantielle Ladung aufweisen. Daraus folgt, daß zwei Faktoren so bestimmt werden sollen, daß für verschiedene Merkmale des Faktors die Ladungen verschwinden, während sie beim anderen Faktor wesentlich sind.

1 Vgl. Jahn, W. / Vahle, H., Die Faktoranalyse, a.a.O., S. 116.

2 Ebenda.

3 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 176.

4 Vgl. Thurstone, L.L., Multiple Factor Analysis, Chicago 1947, S. 112ff und Harman, H.H., Modern Factor Analysis, Chicago 1962, S. 335 zitiert bei Jahn, W. / Vahle, H., Faktoranalyse, a.a.O., S. 113f; ferner Überla, K., Faktorenanalyse, a.a.O., S. 184.

4. Bei mehr als vier Faktoren soll für jedes Spaltenpaar der größte Teil der Merkmale Nullladungen aufweisen.
5. Für jedes Spaltenpaar sollen bei einer geringen Merkmalsanzahl beide Ladungen von Null verschieden sein, d.h., daß die Faktoren so bestimmt werden müssen, daß einige Merkmale bei jedem von zwei Faktoren hoch geladen sind.

Durch eine Faktorrotation unter Beachtung der Prinzipien der Einfachstruktur ist es möglich, neue Faktormuster in den Ausgangsdaten zu entdecken. Schließlich ermöglicht es die Rotation, Hypothesen zu testen, und zwar in dem Sinne, wie gut das Faktormuster bestimmten Daten einer vorher datenmäßig spezifizierten Hypothese genügt.¹

Bei den bisherigen Erörterungen der orthogonalen Rotation wurde davon ausgegangen, daß die zu interpretierenden Faktoren nicht korreliert sind. Hebt man diese Voraussetzung auf, so müssen die Koordinatenachsen des rotierten Koordinatensystems nicht mehr rechtwinklig aufeinander stehen. Während bei der orthogonalen Rotation Lösungstechniken bestehen, deren Validität sich erwiesen hat, ist die Wahl geeigneter schiefwinkliger Rotationsverfahren problematisch.²

Den letzten Schritt bei der Faktorenanalyse bildet die Bestimmung von Faktorwerten. Es sollen hier hypothetische Einflußgrößen ermittelt werden. Während es bei den erörterten Verfahren darum ging, Faktoren zu extrahieren, die die Korrelationsmatrix zu reproduzieren gestatten³, kann man noch einen Schritt weitergehen und versuchen, die Faktorwerte aufgrund der Fundamentalbeziehung der Faktoren-

1 Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 223.

2 Vgl. ebenda, S. 211ff.

3 Vgl. ebenda, S. 235.

analyse (10) zu bestimmen, wenn die Matrix A_{il} bekannt ist. Wenn in (9) E_{lj} aus der Kenntnis von A_{lj} ermittelt wird, ist es möglich, den Beobachtungswerten Faktorwerte zuzuordnen. Da die Berechnung der Matrix F_{lj} aber mit erheblichen Fehlerquellen behaftet ist, wird man in den meisten Fällen eine Schätzung der Faktorenwerte durch multiple Regressionsrechnung vorziehen.¹

¹ Vgl. Überla, K., Faktorenanalyse, a.a.O., S. 241ff.

IV. Die Modellbank im Marketing-Informationssystem

1. Modelle als Komponenten der Modellbank

1.1 Modellarten

Modelle als real-konkrete oder formal-konkrete Abbilder real-konkreter oder real-konzeptueller Systeme¹ lassen sich beispielsweise nach dem Zeitbezug (statisch - dynamisch), nach den in ihnen enthaltenen Variablen (monovariabel - multivariabel), nach der Art des angewandten mathematischen Verfahrens (marginalanalytische Modelle - Modelle der mathematischen Programmierung) und nach einer Vielzahl weiterer - oft nicht überschneidungsfreier Kriterien - einteilen. Für die Zwecke der Typisierung von Marketing-Modellen ist ein generelles Kriterium hier besser geeignet als eine Vielfalt von Klassifizierungsmerkmalen, da dann zumindest eine überschneidungsfreie Einordnung von Modellen ermöglicht wird.

Ein derartiges Einteilungskriterium ist der Modellzweck. Ausgehend vom Modellzweck kann zwischen Prognose- und Entscheidungsmodellen unterschieden werden.² Prognosemodelle ermöglichen die Vorhersage von relevanten Ereignissen in Abhängigkeit von den von der Unternehmung kontrollierbaren und nicht kontrollierbaren Variablen. Bei diesen Modellen handelt es sich um empirisch-kognitive Aussagensysteme in Form von nomologischen Hypothesen, die es ermöglichen, marketingrelevante Phänomene zu erklären. Mittels eines Prognosemodells soll z.B. die zukünftige Absatzentwicklung für einen mehr oder weniger langen Zeitraum vorherbestimmt werden. In der Regel wird man bei diesem Modelltyp so vorgehen, daß man aus den vergangenen Entwicklungen Anhaltspunkte für zukünftige Entwicklungen zu gewinnen trachtet.

1 Vgl. Grochla, E., Modelle und betriebliche Informationssysteme, Arbeitspapier No. 1, Februar 1973, S. 2.

2 Vgl. Meffert, H., Gute Zukunftschancen für mathematische Modelle, in: Wirtschaftswoche, 27. Jg., Nr. 19, 1973, S. 75ff, hier S. 76f.

Zusätzlich zu den aus Vergangenheitsdaten gewonnenen Erkenntnissen versucht man, sicheres oder wahrscheinliches Wissen über die Zukunft zu berücksichtigen.

Die Berücksichtigung sicheren und wahrscheinlichen Wissens führt zur Einteilung von Prognosemodellen in konditionale und nicht-konditionale Prognosemodelle.¹ Konditionale Prognosemodelle zeichnen sich dadurch aus, daß in ihnen die vom Unternehmen kontrollierten Variablen enthalten sind (sicheres Wissen), während in den nicht-konditionalen Prognosemodellen auch Variablen eingehen, die von der Unternehmung nicht kontrolliert werden können (wahrscheinliches Wissen).

Die konditionalen Prognosemodelle zeigen die vom Unternehmen erwarteten Konsequenzen in Hinblick auf spezifische Ziele wie Marktanteil, Umsatzentwicklung und Gewinn bei den von der Unternehmung kontrollierten Variablen auf. In diese Modellklasse lassen sich die meisten "stochastischen Modelle des Käuferverhaltens" einordnen, während die nicht-konditionalen Prognosemodelle mit den statistischen Prognoseverfahren übereinstimmen.

Werden in Prognosemodelle eine oder mehrere Zielfunktionen sowie die Randbedingungen der zu beeinflussenden Situationen und explizite Entscheidungsvariablen aufgenommen, so handelt es sich um Entscheidungsmodelle. Entscheidungsmodelle haben die Aufgabe, die Konsequenzen alternativer Handlungen von Entscheidungsträgern zu bestimmen, die - gemessen an der oder den vorgegebenen Zielfunktionen - optimal sind.²

Ein Versuch, die Leistungsfähigkeit dieser Modelltypen zu beurteilen, führt schließlich - als zusätzliches Einteil-

1 Zur Einteilung zwischen konditionalen und nicht-konditionalen Prognosemodellen vgl. Krautter, J., Marketing-Entscheidungsmodelle, a.a.O., S. 17f.

2 Vgl. Meffert, H., Gute Zukunftschancen für mathematische Modelle, a.a.O., S. 76.

lungskriterium - zur Berücksichtigung der zur Modellösung angewandten Verfahren.

Hierbei kann in Anwendung an ANSOFF¹ zwischen analytischen, heuristischen und quasi-analytischen Modellen unterschieden werden. Diese Modelltypologie beruht darauf, daß Modelle mittels Methoden der deduktiven Problemanalyse überschaubare, oft mathematisch formulierte Teilzusammenhänge ausgliedern.²

Analytische Modelle zeichnen sich dadurch aus, daß für ein der Struktur nach abgrenzbares Entscheidungsproblem mit einer skalaren Zielfunktion ein Lösungsalgorithmus zur Verfügung steht. Damit sind analytische Modelle primär ergebnisorientiert. Nach der Lösungsfindung wird die Suche nach neuen Lösungen abgebrochen. Ausgehend von dieser Lösung fließen die Ergebnisse nicht wieder in den Entscheidungsprozeß ein, sondern werden nach Beendigung dieses Prozesses überprüft, um dann erst wieder neue Informations-Beschaffungsvorgänge zu induzieren. Der sich hier vollziehende Prozeß ist somit sequentieller Natur.³

Im Gegensatz dazu verzichten heuristische Methoden auf eine Logisierung der zu untersuchenden Erscheinungen.⁴ Es wird auf die Strukturierung des zu lösenden Problems und die Anwendung eines Algorithmus verzichtet. Eine generelle Methodologie, die sich auf sämtliche heuristisch zu lösenden Probleme anwenden läßt, ist nicht vorhanden. Nach

1 Vgl. Ansoff, H.I., A Quasi-Analytic Method for Long Range Planning, in: Alexis, M. / Wilson, C.Z., Hrsg., Organizational Decision Making, Englewood Cliffs, N.J. 1967, S. 427ff, hier S. 431ff.

2 Vgl. Meffert, H., Zum Problem der betriebswirtschaftlichen Flexibilität, a.a.O., S. 799.

3 Vgl. derselbe, Betriebswirtschaftliche Kosteninformationen, Wiesbaden 1968, S. 86ff.

4 Vgl. dazu derselbe, Zum Problem der betriebswirtschaftlichen Flexibilität, a.a.O., S. 800.

ANSOFF¹ besteht das Hauptcharakteristikum der heuristischen Methode darin, daß sich der betreffende Entscheidungsträger auf seine Erfahrungen hinsichtlich der Problemlösung verläßt und nicht in isolierender Abstraktion versucht, einen Lösungsalgorithmus zu finden.

Im Marketing gibt es eine Reihe von Problemen, die sich einer Lösung mittels analytischer Methoden entziehen, wie z.B. bestimmte Fragestellungen aus der Produkt- und Diversifizierungspolitik und viele Allokationsprobleme, die sich insbesondere beim Marketing-Mix stellen. Aufgrund teilweise schlechter Strukturierung derartiger Entscheidungsprobleme läßt sich kein Algorithmus finden, der in toto zur Lösung der gestellten Probleme beiträgt, während weiterhin eine Problemlösung nicht allein auf heuristische Verfahren zurückgreifen braucht, da sich durchaus bestimmte Unterprobleme mittels Algorithmen lösen lassen. Um derartige Probleme zu lösen, ist von ANSOFF das quasi-analytische Lösungsverfahren konzipiert worden, das sich sowohl bestimmter Gedanken der analytischen als auch der heuristischen Verfahren bedient und somit eine Mittelstellung zwischen beiden einnimmt.²

Die mittels den quasi-analytischen Verfahren ermittelten Lösungen werden nicht als endgültig angesehen, sondern als vorläufige Lösungen in die verschiedenen Stadien des Entscheidungsprozesses zurückgeführt.³ Während des Lösungsprozesses treten Rückkoppelungen zwischen Analyse und Problemformulierung ein, die dazu führen, das Problem neu zu formulieren.⁴ Dieser Sachverhalt ist darauf zurückzuführen, daß bei quasi-analytischen Verfahren keine Theorie zur Verfügung steht, so daß es während des Problemlösungs-

1 Vgl. Ansoff, H.I., A Quasi-Analytic Method for Long Range Planning, a.a.O., S. 430.

2 Vgl. ebenda, S. 431.

3 Vgl. ebenda, S. 434f.

4 Vgl. ebenda, S. 434.

prozesses generierter Informationen bedarf, um die ursprünglich aufgestellten Hypothesen zu modifizieren.

Insgesamt kann der Unterschied zwischen analytischen und quasi-analytischen Modellen darin gesehen werden "as a difference in the kind of compromise in making an abstraction from the real world. OR compromises in favor of precision and mathematical completeness, the QA in favor of comprehensiveness and accurate description of reality."¹

Durch Strukturierung der Vorgehensweise des Entscheidungsträgers weist das quasi-analytische Verfahren einen fundamentalen Vorteil gegenüber heuristischen Verfahren auf, das ihn dazu nötigt, sein Lösungsvorgehen explizit zu machen und die Lösung zu bewerten.² Darüber hinaus wird deutlich, daß Unterprobleme beim quasi-analytischen Vorgehen algorithmisch oder heuristisch gelöst werden können. Die von ANSOFF stammende Charakteristik der analytischen, quasi-analytischen und heuristischen Vorgehensweise ist in folgender Tabelle zusammengestellt:³

1 Vgl. Ansoff, H.I., A Quasi-Analytic Method for Long Range Planning, a.a.O., S. 436f.

2 Vgl. ebenda, S. 434.

3 Diese Tabelle ist in Anlehnung an Ansoff, H.I., A Quasi-Analytic Method for Long Range Planning, a.a.O., S. 436 und Meffert, H., Die Flexibilität in betriebswirtschaftlichen Entscheidungen, Habilitationsschrift, München 1968, S. 378 (unveröffentlicht) entwickelt worden.

Verfahren Merkmal	analytisch	quasi-analytisch	heuristisch
Modellansatz	ergebnis-orientiert	prozeßorientiert	prozeßorientiert
Abstraktionsgrad	hoch	mittel	gering
Sprache	numerisch/symbolisch	numerisch/logisch	logisch
Modellformulierung	explizit	implizit	keine
Algorithmus	vorhanden	nur bei Unterproblemen vorhanden	fehlt
Analyse	sequentiell	rückgekoppelt	ad libitum
Zielfunktion	skalar	vektoriell	fehlt
Berücksichtigung qualitativer Faktoren	Präferenzordnungen, Indifferenzkurven	Bewertung	ad libitum

Tab. 10: Charakteristik von analytischen, quasi-analytischen und heuristischen Modellen

Die Hauptcharakteristiken der quasi-analytischen Methode können somit wie folgt zusammengefaßt werden: Die Methode ist aufgrund der schlechten Strukturierung des Entscheidungsprozesses mehr prozeß- als zielorientiert.¹ Der Problemrahmen ist qualitativer und quantitativer Natur, wobei

¹ So stellt Harnischfeger fest: "Die Ziele stellen ... eine Operationalisierung und eine Interpretation der Zwecke dar. Sie sind Interpretationen, weil sich in den Zielen erst darstellt, was im einzelnen mit den Zwecken gemeint ist, die wegen der notwendig allgemeinen Formulierung vielfältiger Auslegung zugänglich sind." Harnischfeger, H., Planung in der Sozialstaatlichen Demokratie, Berlin 1969, S. 93.

Unterprobleme quantitativ gelöst werden können. Die allgemeine Lösungsmethodik ist adaptiv, wobei Rückkopplungen zwischen den jeweils erzielten Lösungen und Problemformulierungen bestehen. Nach der Erläuterung der Modellarten soll im folgenden untersucht werden, worin die Funktionen und der Inhalt der Modellbank als einer zentralen Komponente eines Marketing-Informationssystems bestehen.

1.2 Funktionen und Inhalt der Modellbank

Die Modellbank eines Marketing-Informationssystems läßt sich neben der Daten- und Methodenbank als dritte bestimmende Komponente dieses Systems definieren. Sie ist die Zusammenfassung sämtlicher programmierter Marketing-Modelle.¹ Der Vorteil dieser Programmpakete wird besonders durch das "Zusammenspiel" mit der Datenbank deutlich: Unter der Prämisse eines hohen Datenvolumens werden die Programme mit Daten versorgt, die in der Datenbank des Marketing-Informationssystems gespeichert sind. Aufgrund der komplexen Entscheidungsproblematik im Marketing bieten diese Programme die Möglichkeit, modellmäßig faßbare Sachverhalte sehr schnell zu berechnen, so daß die Ergebnisse den Entscheidungsträgern als Entscheidungshilfen zur Verfügung gestellt werden können.

Für das Vordringen der Modellanalyse im Marketing² lassen sich eine Reihe von Gründen anführen. Der empirische Grund ist die Diskrepanz zwischen Beobachtung und theoretischer Vorhersage, die sich immer dann ergibt, wenn der "Beobachtungsradius" infolge neuer informationsgewinnender und -verarbeitender Techniken verbessert wird. Darüber hinaus führen Modelle, wenn ausreichende Homorphie zwischen Modell und Realität besteht, dazu, daß sie - wie z.B. Simulationsmodelle - es erst ermöglichen, eine Reihe bis dahin nicht vorzunehmender Beobachtungen zu machen. In diesem Zusammenhang werden die Vorteile der Modellanalyse recht deutlich. Sie gestattet eine bessere gedankliche Durchdringung von Phänomenen, ein erhöhtes Verständnis von Struktur und Dynamik von Marketing-Prozessen und die explizite Zielformulierung. Diese Vorteile versucht man auch im Bereich

1 Zur mathematischen Modellanalyse vgl. z.B. Meffert, H. und Mitarbeiter, Die Anwendung mathematischer Modelle im Marketing, a.a.O.; Krautter, J., Marketing-Entscheidungsmodelle, a.a.O. und die dort angeführte Literatur.

2 Vgl. Koxholt, R., Die Simulation - Ein Hilfsmittel der Unternehmensforschung, München - Wien 1967, S. 15ff.

der Informationssysteme zu nutzen, indem Verarbeitungsprogramme entwickelt wurden, die es letztlich gestatten sollen, im Marketing bessere Entscheidungen zu fällen.

Ebenso wie bei der Datenbank ist es bei der Modellbank notwendig, die einleitend angeführte Definition dahingehend zu präzisieren und zu untersuchen, inwieweit eine spezielle Modellbank für das Marketing konzipiert werden kann. Da die in der Modellbank enthaltenen Modelle mit den Daten aus der Datenbank implementiert werden, besteht zwischen beiden Subsystemen eines Marketing-Informationssystems eine enge Beziehung.

Unter Berücksichtigung der Tatsache, daß es nicht möglich ist, sämtliche Modelle zu analysieren, die für die Aufnahme in die Modellbank eines Marketing-Informationssystems geeignet sind, bietet sich die Untersuchung von Modelltypen für spezifische Marketingfunktionen an. Eine Möglichkeit zur Modellkategorien-Bildung besteht darin, von den Modellfunktionen im Marketing auszugehen.

Die Funktion einer Modellbank kann nur aufgrund der Menge der Funktionen von den in ihr enthaltenen Marketing-Modellen untersucht werden. "Die Konstruktion und Verwendung von Modellen erfolgt stets zu bestimmten erkenntnistheoretischen und gestaltungsspezifischen Zwecken."¹ Es muß also davon ausgegangen werden, daß ein Modell nur dann von einem Benutzer gewählt wird, wenn es ihm auf diese Weise gelingt, spezifische Aspekte und Zusammenhänge im pragmatischen Sinn besser erklären oder gestalten zu können. Die Erreichung dieses Ziels setzt voraus, daß die mit der Prozeßvornahme der zu berücksichtigenden Marketingfunktionen verbundenen Prozesse der Gewinnung und entscheidungslogischen Verarbeitung von Informationen im Sinn eines Modells formalisiert werden können. Geht man von der funktionalen

1 Grochla, E., Modelle und betriebliche Informationssysteme, a.a.O., S. 3.

Strukturierung des Marketing aus, so stehen insbesondere die folgenden Fragestellungen im Vordergrund, deren Beantwortung durch Marketing-Modelle gewährleistet sein soll:

(1) Marketing-Logistik¹

- Standortwahl: Wo liegen die "gewinnoptimalen" Standorte für Produktionsstätten und Auslieferungsläger? Beantwortung durch Standortmodelle.
- Lagerpolitik: Wie können Lagerbewegungen und Bestände von Fertigprodukten optimiert werden? Beantwortung durch Lagerhaltungsmodelle.
- Transport- und Lieferzeitenpolitik: In welcher Reihenfolge sind die Auslieferungsläger und/oder Kunden zu beliefern? Beantwortung durch Travelling-Salesman-Modelle. Wie kann eine kostenminimale Verteilung der Produkte von den Produktionsstätten an die Auslieferungsläger vorgenommen werden? Beantwortung durch Allokationsmodelle.

(2) Kontrahierung

- Preispolitik: Wie soll ein Produktpreis bei Einführung eines neuen Produkts, bei der Preisänderung von Konkurrenzprodukten gesetzt werden? Beantwortung durch preispolitische Modelle.
- Rabattpolitik: Wie sollen optimale Rabattstaffeln festgelegt werden? Beantwortung durch Rabatt-Simulations-Modelle.
- Zahlungsbedingungen und Konditionenpolitik: In welcher Höhe sollen Lieferantenkredite, Ausstattungskredite, Boni und Skonti gewährt werden? Beantwortung durch Entscheidungsmodelle.

(3) Verkäuferereinsatz (Sales Force)

- Tourenplanung: In welcher Reihenfolge sind die Kunden zu besuchen? Beantwortung durch Travelling-Salesman-Modelle.
- Besuchsisintensität: In welcher Frequenz sind Kunden zu besuchen? Beantwortung durch Travelling-Salesman-Modelle.
- Absatzbezirke: Welche Absatzbezirke sind zu bilden; wie oft sollen diese Absatzbezirke durch Vertreter bereist werden? Beantwortung durch Allokationsmodelle.

¹ Vgl. dazu Zangemeister, Ch., Marketing-Informationssysteme: EDV-Einsatz im gesamten Absatzbereich, Bd. 5 der SCS Schriftenreihe, Hamburg o.J., S. 27ff.

(4) Absatzprogramm bzw. Sortiment:

Welche Mengen können von einem Produkt oder von einem Sortimentsteil abgesetzt werden? Beantwortung durch Prognose- und Sortimentsmodelle.

(5) Werbung und Verkaufsförderung

- Budgetaufstellung: Wie hoch ist das Werbebudget für einzelne Produkte festzusetzen? Beantwortung durch Werbebudgetierungsmodelle.
- Medienselektion: Welche Werbemittel sind wann, wie häufig und in welcher Form einzusetzen? Beantwortung durch Mediaselektionsmodelle.

(6) Produktinnovation

- F + E-Budget: Wie soll das Forschungs- und Entwicklungsbudget auf einzelne neu zu entwickelnde Produkte verteilt werden? Beantwortung durch Budgetierungsmodelle.
- Ideenbewertung: Wie und nach welchen Kriterien sollen Neuprodukte hinsichtlich ihres Markterfolgs bewertet werden? Beantwortung durch Bewertungsverfahren.
- Produkteinführung: Wann soll ein neu entwickeltes Produkt am Markt eingeführt werden? Beantwortung durch Modelle zur Produktneueinführung.

(7) Wahl der Absatzkanäle (Chanel Design)

Über welche Absatzkanäle soll ein Produkt abgesetzt werden? Beantwortung durch Bewertungsmodelle.

(8) Gesamt-Mix

Wie ist das optimale Marketing-Mix für sämtliche Produkte zu bestimmen und festzulegen? Beantwortung durch produktbezogene Gesamt-Mix-Modelle.

Dieser keineswegs vollständige Überblick über die im Rahmen von Marketingfunktionen anwendbaren Modelle läßt sich noch dahingehend erweitern, daß eine subjektive Einstufung von Problemlösungsmethoden für einzelne Marketingfunktionen durch Relevanzziffern von 1 bis 10 vorgenommen wird. Die folgende Tabelle¹ ordnet den oben angeführten Marketing-Entscheidungsproblemen diesbezügliche Relevanzziffern zu.

¹ In Anlehnung an Zangemeister, Ch., Marketing Informationssysteme, a.a.O., S. 28.

Lösungsverfahren	Verfahren der mathematischen Problemlösung	Infinite-dimensionalrechnung	Multivariate statistische Verfahren							Graphentheoretische Verfahren und Netzplantechnik	Markov-Verfahren	Entscheidungsbaumverfahren	Nutzwertanalyse	Spieltheorie	Simulation	heuristische Verfahren							
			lineare Programmierung	ganzzahlige Programmierung	dynamische Programmierung	wohlstrukturiert (routinemäßig)																	
						operative Systeme																	
Marketing-Entscheidungsprobleme	Strukturierungsgrad der Entscheidungsprobleme																						
	Marketing-Logistik																						
	- Standortwahl																						
	- Lagerpolitik																						
	- Transport- und Lieferzeitenpolitik																						
	Kontrahierung																						
	- Preispolitik																						
	- Rabattpolitik																						
	- Zahlungsbedingungen und - Konditionenpolitik																						
	Verkaufereinsatz (Sales Force)																						
	- Tourenplanung																						
	- Besuchintensität																						
	- Absatzbezirke																						
	Absatzprogramm bzw. Sortiment																						
	Werbung und Verkaufsförderung																						
	- Budgetaufstellung																						
	- Medienselektion																						
	Produktinnovation																						
	- P + E-Budget																						
	- Ideenbewertung																						
	- Produktneinführung																						
	Wahl der Absatzkanäle (Channel Design)																						
	Gesamt-Mix																						

Tab. 11: Einstufung von Problemlösungsmethoden durch Televanziffern von 1 bis 10

Als Ergebnis dieser Untersuchung ergibt sich folgendes:¹
Bei der Marketing-Logistik dominieren Verfahren der mathematischen Programmierung und heuristische Verfahren.² Beim Kontrahierungs-Mix stehen sowohl Verfahren der Infinitesimalrechnung als auch Simulation und heuristische Verfahren im Vordergrund. Der Verkäufereinsatz ist nach subjektiver Einstufung vor allem durch mathematisch-statistische Verfahren, Entscheidungsbaumverfahren, Simulation und heuristische Verfahren zu lösen, während beim Absatzprogramm bzw. Sortiment die Verfahren der mathematischen Programmierung dominieren. Werbung und Verkaufsförderung sind vor allem durch Verfahren der mathematischen Programmierung, durch multivariate statistische Verfahren, Simulation und heuristische Verfahren einer Lösung näher zu bringen. Bei der Produktinnovation scheinen Infinitesimalrechnung, multivariate statistische Verfahren, Netzplantechnik und Entscheidungsbaumverfahren von größter Bedeutung zu sein, während bei der Wahl der Absatzkanäle Simulation und heuristische Verfahren im Vordergrund stehen. Schließlich sind beim Gesamt-Mix Simulation und heuristische Verfahren die erfolgversprechendsten Methoden.

Das Ergebnis dieser Untersuchung ist weiterhin darin zu sehen, daß mit zunehmendem Strukturierungsgrad der Entscheidungsprobleme die Anzahl wohlstrukturierter Entscheidungen, die durch operative Marketing-Systeme gefällt werden können, abnimmt. Gleichermäßen nehmen die Bereiche adaptiver (administratives System) und schlechtstrukturierter Entscheidungen (politisches System) zu. Während bei den wohlstrukturierten Entscheidungen die Verfahren der mathematischen Programmierung vorherrschen, dominieren bei den schlechtstrukturierten Entscheidungen die Simulation und heuristische Verfahren.

1 Vgl. dazu Tab. 10.

2 Der Begriff "heuristisch" wird in diesem Zusammenhang als planmäßiger Suchprozeß nach neuen Lösungen verstanden, während die "heuristische Methode" im Sinne Ansoffs primär auf die Erfahrung abstellt.

Die Selektion und Entwicklung geeigneter Marketing-Modelle kann aber nur dann realisiert werden, wenn es gelingt, Kriterien zu entwickeln, die über die Auswahl und Entwicklung dieser Modelle Auskunft geben.

2. Ansätze zur Modellselektion und -entwicklung

2.1 Der theoretisch-intersubjektive Ansatz

Die Wissenschaftstheorie geht bei der Beschreibung individueller Phänomene von allgemeinen Gesetzhypothesen aus, mit deren Hilfe empirische Phänomene erklärt und prognostiziert werden können.¹ Es wird dabei zwischen Aussagen unterschieden, deren Geltung allein logisch begründet ist (analytische Aussagen) und Aussagen, die zwar in sich logisch sind, aber zumindest auf einer empirischen Relation beruhen (synthetische Aussagen).² Die Zusammenfassung von analytischen und synthetischen Aussagen bietet die Möglichkeit, theoretische Grundlagen über die Aufnahme von Marketing-Modellen in die Modellbank eines Marketing-Informationssystems zu entwickeln.

Es soll zunächst erläutert werden, was unter einer Theorie zu verstehen ist und worin die Beziehung zwischen dieser Theorie und empirisch beobachtbaren Phänomenen besteht. Es wurde soeben bereits eine implizite Definition einer Theorie gegeben: Es handelt sich um einen Erklärungsversuch einer Reihe von empirisch beobachtbaren Phänomenen mittels Modellen.³ Die Aufgabe der Theorie besteht darin, ein Modell zu finden, das - auf empirische Situationen angewendet - die gleichen Daten liefert wie diese. Eine Theorie ist um so besser, je mehr Daten sie zu deuten vermag und je genauer sie die Ergebnisse künftiger Beobachtungen voraussagen gestattet. Wesentliches Beurteilungskriterium eines Modells ist somit dessen prognostische Relevanz.⁴

1 Vgl. z.B. Popper, K.R., Logik der Forschung, 3. Aufl., Tübingen 1969, S. 199ff.

2 Vgl. Albert, H., Probleme der Theoriebildung, in: Albert, H., Hrsg., Theorie und Realität, Tübingen 1964, S. 3ff, hier insbesondere S. 20, 24, 31 und 50.

3 Vgl. dazu Katterle, S., Normative und explikative Betriebswirtschaftslehre, Göttingen 1964, S. 122f.

4 Vgl. Stegmüller, W., Hauptströmungen der Gegenwartsphilosophie - Eine kritische Einführung, Stuttgart 1960, S. 456.

Wesensbestimmendes Merkmal einer Theorie ist, daß sie beides, nämlich Vorhersage und Interpretation zuläßt. Eine Theorie, die verfügbare Ergebnisse nicht zu erklären vermag, ist offensichtlich unbrauchbar.

Darüber hinaus sind oft mehrere Theorien über einen Erkenntnisbereich aufstellbar, die sämtliche bekannten Daten in befriedigender Weise wiedergeben. Die einzige Möglichkeit, zwischen mehreren Theorien zu entscheiden, besteht im Ersinnen eines Experiments bzw. eines Tests, für die die Theorien unterschiedliche Ergebnisse vorhersagen. Nach diesen Überlegungen ist es einsichtig, daß keine Theorie jemals "endgültig" sein kann, da man stets mit der Wahrscheinlichkeit rechnen muß, daß künftige Daten nicht mit den theoretischen auf Basis des Modells gewonnenen Vorhersagen übereinstimmen.

Ferner ist einleuchtend, daß eine neue, bessere Theorie oft durch Verbesserung bzw. Abänderung einer schon existierenden Theorie gefunden werden kann. Selbst dann, wenn eine neue Theorie völlig anders als ihre "Vorgänger" ist, muß sie in jenen Bereichen, in denen die alten Theorien heuristische Kraft besaßen, die durch empirische Untersuchungen festgestellt werden konnte, die gleichen Ergebnisse wie diese liefern. Folglich muß sie die älteren Theorien in einem entsprechenden Ausmaß umfassen.

Angesichts der Komplexität derartiger Theorien ist eine intersubjektive Nachprüfbarkeit miteinander verbundener nologischer Hypothesen (Satzsysteme) daran gebunden, daß gewisse Vorgänge reproduzierbar bzw. wiederholbar gemacht werden können.¹ Im Bereich der Naturwissenschaften bedient man sich dazu des Experiments, während im Marketing Modelle

¹ Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 19.

als Abbild der Realität oft an die Stelle des Experiments treten.¹

Zentrales Anliegen dieses theoretisch-intersubjektiven Ansatzes zur Modellselektion und -entwicklung ist die Diskussion der Isomorphie- bzw. Homomorphieproblematik. Bei isomorpher bzw. homomorpher Abbildung der Realität ist es aufgrund der Modellaussagen möglich, zu einer intersubjektiven Nachprüfbarkeit zwischen Modell und Realität zu gelangen.

Ein Modell wird dann als isomorph bezeichnet, wenn es Aussagen liefert, die sowohl für das Modell selbst als auch für das Realsystem zutreffen und umgekehrt.² Man kann zwar nicht sagen, daß Modell und Realität identisch sind, da beides als unvergleichbar nicht miteinander verglichen werden darf. So besteht keineswegs Identität zwischen einer Unternehmung, die auf einem bestimmten Markt Angebotsmonopolist ist und dem Modell des Angebotsmonopols. Der hier definierte Begriff der Isomorphie ist ein Spezialfall der Homomorphie. Allgemein versteht man unter Homomorphie eines Modells die Abbildung eines Systems auf ein anderes System, wobei gewisse Struktureigenschaften erhalten bleiben.³

Um den Begriff der Homomorphie präziser zu fassen, benötigt man den Abbildungsbegriff. "Eine Abbildung f aus einer Menge

1 Vgl. Pagés, R., Das Experiment in der Soziologie, in: König, R., Hrg., Handbuch der empirischen Sozialforschung, Bd. I, Stuttgart 1962, S. 415ff, hier S. 427f. Dies gilt insbesondere für die "deterministischen Experimente" (Pagés), die in ihrer formalen Struktur große Ähnlichkeiten mit den deterministischen Modellen der mathematischen Modellanalyse aufweisen.

2 Vgl. zum Isomorphiebegriff Bertalanffy, L., General Systems Theory, in: Human Biology 1951, S. 303ff. Isomorphie ist ein Begriff, der anzeigen soll, daß miteinander verglichene Systeme aus identischen Elementen bestehen.

3 Vgl. dazu Craik, K., The Nature of Explanation, a.a.O., S. 54 und Lenz, H., Grundlagen der Elementarmathematik, Berlin 1960, S. 60.

M in die Menge M' ist dann eine Teilmenge der Menge $M \times M'$ aller geordneten Paare $[a, a']$ mit $a \in M$ und $a' \in M'$. Kommt dabei jedes Element von M mindestens einmal als erstes Element in einem Paar aus f vor, so spricht man von einer Abbildung $v o n M$, kommt jedes Element von M' mindestens einmal als zweites Element in einem Paar aus f vor, so spricht man von einer Abbildung $a u f M'$. Kommt jedes Element von M höchstens einmal als erstes Element in einem Paar aus f vor, so heißt die Abbildung eindeutig oder eine Funktion."¹

Dieser universelle Abbildungsbegriff findet in gewissen vorzunehmenden Transformationen seine Konkretisierung. Unter Transformation versteht man die Befolgung einer Zuordnungsvorschrift, welche einer Funktion oder einer Menge eine andere Funktion oder Menge zuordnet.² Ohne im Detail auf den recht schwierigen mathematischen Hintergrund dieses Problemkreises einzugehen, ist hier zu betonen, daß sich Transformationen mit mathematischen Abbildungsproblemen befassen, wobei die Zuordnung zwischen Original- (Realität) und Bildbereich (hier Modell) im wesentlichen umkehrbar eindeutig ist.³ Als Beispiele für in ökonomischen Bereichen angewendete Transformationen lassen sich die z -Transformation der Statistik und die Laplace-Transformation der Regelungstheorie anführen.⁴

Ziel jeglicher Transformation ist die "Übersetzung" eines mathematischen Problems - z.B. die Lösung einer Differen-

1 Kleine Enzyklopädie der Mathematik, Leipzig 1965, zitiert nach Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 270.

2 Vgl. Behnke, H. et al., Hrsg., Mathematik, Bd. 1, a.a.O., S. 268.

3 Vgl. Landgraf, C. / Schneider, G., Elemente der Regelungstechnik, Berlin-Heidelberg-New York 1970, S. 53.

4 Zur z -Transformation vgl. z.B. Fisher, R.A., Statistische Methoden für die Wissenschaft, 12. Aufl., Edinburgh-London 1956 und zur Laplace-Transformation Holbrook, J.G., Laplace-Transformationen, Braunschweig 1970.

tialgleichung - in das entsprechende Problem des Bildbereiches, um es dort lösen und danach die Lösung in den Originalbereich "zurückzuübersetzen".

Dieser Prozeß läßt sich schematisch wie folgt kennzeichnen:¹

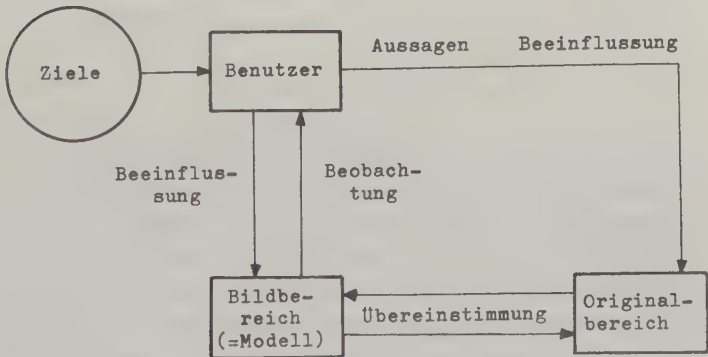


Abb. 15: Zusammenhang zwischen abzubildendem Original, Benutzer und Modell

Der Benutzer eines Modells hat verschiedene Ziele, die darin bestehen, den Originalbereich zu erklären oder zu beeinflussen. Dieses erreicht er dadurch, indem der Originalbereich möglichst "stark" im Bildbereich abgebildet wird. Der Bildbereich kann bei homomorpher Abbildung in den Originalbereich rückübersetzt werden. Stimmen Bild- und Originalbereich nicht überein, hat der Benutzer unter Beachtung des Modells die Möglichkeit, den Original- als auch den Bildbereich dahingehend zu beeinflussen, daß eine möglichst großer Übereinstimmung zwischen Original- und Bildbereich gegeben ist.

¹ Quelle: Grochla, E., Modelle und betriebswirtschaftliche Informationssysteme, a.a.O., S. 4. (Die Nomenklatur wurde teilweise vom Verfasser geändert)

Das Transformationsproblem betrifft dabei die Frage, ob sich das Problem im Originalbereich gegenüber der Transformation als invariant erweist oder nicht. Es betrifft die rein mathematische Seite des Homomorphieproblems, und zwar dergestalt, daß nach der Art der zulässigen Transformationen gefragt wird. Auf einen sehr einfachen Nenner gebracht, lautet die Frage: Welche Transformationen dürfen vorgenommen werden, damit nicht gegen die Homomorphie des Modells verstoßen wird? Zunächst sollen nun die Transformationsarten behandelt werden, die sich bei Modellen feststellen lassen.

Bei der Transformation durch Permutation¹ wird ausschließlich die Reihenfolge der Elemente geändert. Modelle, die sich auf geschlossene Transformationen zurückführen lassen - zu denen ebenfalls die Transformation durch Permutation zählt - zeichnen sich dadurch aus, daß eine Optimallösung letztlich nur durch Umstrukturierung bekannter Daten erreicht wird und keine neuen Datenbeschaffungsaktivitäten vorgenommen werden müssen, um das Problem zu lösen. Derartige Transformationen finden sich z.B. in Zuweisungsmodellen. Als Beispiele aus dem Marketingbereich seien hier die Modelle zur Optimierung des Rundreiseweges ("Vertreterereinsatzmodelle") genannt. Die hier vorgenommenen permutierten Transformationen sind ein Grund dafür, daß sich diese Modellkategorie mittels eines Zuweisungsalgorithmus lösen läßt.

Die Transformation durch Substitution² zeichnet sich dadurch aus, daß hier einige Elemente durch andere ersetzt werden. Die ersetzenden Elemente weisen gegenüber denen der Ausgangssituation völlig unterschiedliche qualitative und quantitative Merkmalsausprägungen auf. Ein äußerst komplexes Problem kann dadurch erheblich vereinfacht und

1 In Anlehnung an Flechtner, H.-J., Grundbegriffe der Kybernetik, a.a.O., S. 355.

2 Vgl. dazu ebenda, S. 354.

einer Lösung zugänglich gemacht werden, wenn es gelingt, Variablen zu eliminieren. Dieses Problem stellt sich in der Statistik immer dann, wenn bestimmte Variablen in Multikollinearität zueinander stehen. Hier wird man die Variablen sukzessive vermindern, solange sich das Ergebnis nicht signifikant verschlechtert. Darüber hinaus besteht die Möglichkeit, nicht kontrollierbare Variablen zu eliminieren. Diese Eliminierung kann einerseits durch völlige Vernachlässigung, andererseits durch Aggregation von Variablen geschehen. Transformation durch Substitution ist für Modelle typisch, die im Lauf der Zeit weiterentwickelt worden sind, wobei ausgehend von einem Grundmodell, bestimmte Variablen, Parameter und/oder funktionale Beziehungen ersetzt wurden. Als Beispiel kann ein im Institut für Marketing entwickeltes Rabattmodell genannt werden, das auf dem Grundmodell rationalen Bestellverhaltens aufbaut und so transformiert wurde, daß die Bestimmung des Einigungsbereiches von für den Verkäufer und Käufer vorteilhaften Rabattsätzen mittels Simulation möglich ist.

Bei der Transformation durch Metamorphose¹ ändert sich ein System in seiner Gesamtheit. Das Modell weist im Zeitpunkt t_0 eine von t_1 völlig abweichende Struktur auf. Diese ist darauf zurückzuführen, daß die das System charakterisierenden Elemente in t_0 mit denen in t_1 nicht mehr identisch sind. Beispiele für Transformationen durch Metamorphose bilden Modellneuentwicklungen, die sich von bisher bekannten Modellen entweder durch die Berücksichtigung eines neuen Problemkreises oder durch bislang unbekannte Algorithmen grundsätzlich bzw. innovativ unterscheiden.

Bei der Transformation durch Elementänderung² bleiben die Elemente zwar dieselben, sie unterscheiden sich aber von der Ausgangssituation durch geänderte Merkmalsausprägungen.

1 Vgl. Flechtner, H.-J., Grundbegriffe der Kybernetik, a.a.O., S. 354.

2 Vgl. ebenda, S. 354f.

Beispiele für Transformationen durch Elementänderung bieten in erster Linie dynamische Modelle, bei denen entsprechend der zeitlich vertikalen Datenentwicklung eine Aktualisierung von Variablen und Parametern vorgenommen werden muß. So können z.B. die Charakteristiken von Variablen und Parametern geändert werden, um ihre Handhabung im Modell zu erleichtern. Erweist sich die Forderung nach Stetigkeit der Datenentwicklung als problematisch, müssen von Fall zu Fall beobachtete Werte in die Funktionen des Modells eingesetzt werden. Umgekehrt kann auch Fluidisierung erfolgen, indem eine diskrete durch eine kontinuierliche Variable ersetzt wird. Darüber hinaus kann die Variabilität sehr gering sein, so daß aus Vereinfachungsgründen die Variable im Modell als Konstante angesehen wird. Weiterhin besteht die Möglichkeit, mit ganzzahligen und nicht-ganzzahligen Ausprägungen von Variablen zu arbeiten, um bestimmten vorher festgelegten Rahmenbedingungen einer Lösung Rechnung zu tragen.

Eine grundlegende Möglichkeit, ein mathematisches Modell zu vereinfachen, besteht darin, entweder die funktionale Form des Gesamtmodells oder einzelne seiner Komponenten zu ändern. Beispielsweise ist es bei a-priori festliegenden Lösungsbereichen sinnvoll, nichtlineare Funktionen durch lineare zu substituieren.

Auf der konkret-rechnerischen Seite tritt das Transformationsproblem in der Gestalt auf, unterschiedliche Größen durch Skalierung komensurabel zu machen.¹ Es bedarf eines Transformationsprozesses, um eine Skala in eine andere oder mehrere Skalen in eine gemeinsame übergeordnete zu überführen. Beispiele bieten vor allem die in der Marktfor-

1 Zu den Skalierungsproblemen vgl. z.B. Osgood, C.E. / Suci, G.J. / Tannenbaum, P.H., The Measurement of Meaning, Urbana, III. 1957; Osgood, C.E., The Nature and Measurement of Meaning, in: Psychological Bulletin No. 49, 1952, S. 197ff; Spiegel, B., Die Struktur der Meinungsverteilung im sozialen Feld, Bern - Stuttgart 1961, S. 41ff.

schung angewandten Skalierungstechniken. Insbesondere lassen sich Rangreihen nur dann bilden, wenn sich die individuellen Rangreihen von Testpersonen durch Transformation in eine gemeinsame übergeordnete Rangreihe überführen lassen.

Bei der Homomorphie von Modellen und Systemen lassen sich drei Homomorphiegrade unterscheiden, die als schwache, starke und strenge Homomorphie bezeichnet werden sollen.¹ Für die schwache Homomorphie wird verlangt, daß die Struktureigenschaften des Systems sich in der Abbildung feststellen lassen. Bezeichnet man mit M die Elementmenge des Systems, mit S die Systemstruktur, d.h. eine Folge von S Relationen über M , so ist bei schwacher Homomorphie das geordnete Paar $[M, S]$ hinsichtlich seiner Struktureigenschaften mit der Abbildung $[M', S']$ identisch. Bei der schwachen Homomorphie kann es aber sein, daß eine Relation aus der Folge von S Relationen über M für das Original nicht zutrifft, während die entsprechende Relation für die Abbildung M' gilt. Beispielsweise gilt dieses für ein Marketing-Modell, das erweiterungsfähig sein soll und aufgrund dieses Ziels eine Relation enthält, die im Original nicht enthalten ist.²

Bei starker Homomorphie ist dieses nicht möglich. Hier gilt, daß die Relationen über M sämtlich auch über M' zutreffen. Die Zwischenstellung der starken Homomorphie ist dadurch gekennzeichnet, daß zumindest ein entsprechendes Relationenpaar sowohl im Modell als auch in

¹ Die folgenden Ausführungen lehnen sich an Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 269-274 an.

² Vgl. dazu den Punkt V./3.5 dieser Arbeit.

der Realität festzustellen ist.¹ Bezeichnet man mit M wiederum die Elementmenge, mit R eine zweistellige Relation über M , so daß $S = [R_1, R_2]$ gilt, so läßt sich ein System, das aus drei Elementen besteht, in der Form $[M, R] = \{ \{A, B, C\}, \{ [A, B], [B, C] \} \}$ darstellen. Die zweistellige Relation bringt zum Ausdruck, daß jeweils zwei Elemente in Relation zueinander stehen. Der oben angeführte Ausdruck $[M, R]$ ist die formalisierte Schreibweise einer Beziehung, wie sie sich in einem Distributionssystem feststellen läßt. Die Größen A , B und C können z.B. bestimmte Elemente dieses Distributionssystems sein. Wird z.B. angenommen, daß A und B Großhändler sind und C ein Einzelhändler ist, so bestehen zwischen den Großhändlern A und B und zwischen dem Großhändler B und dem Einzelhändler C bestimmte Beziehungen, z.B. Kommunikationsbeziehungen.

Zur Unterscheidung vom Realsystem ist es möglich, A , B und C durch A' , B' und C' zu ersetzen, um zu verdeutlichen, daß A ein Großhändler, A' z.B. ein Element in einem Distributionssystem ist. Bilden z.B. der Großhändler B und der Einzelhändler C eine freiwillige Kette, so können sie als Einheit betrachtet und mit D bezeichnet werden. Es ergibt sich dann das neue System $[M'', R''] = \{ \{A, D\}, \{ [A, D], [D, D] \} \}$. In diesem Fall ist die Homomorphie stark, da zu $[A, D] \in R''$ und $[D, D] \in R''$ es jeweils ein entsprechendes Paar in R gibt, nämlich $[A, B] \in R$ und $[B, C] \in R$.

Aus diesen Ausführungen ist ersichtlich, daß das Modell eines Systems im allgemeinen weder homomorph noch isomorph ist. "Sowohl im Realsystem als auch im Modell gibt es Ele-

¹ Starke Homomorphie zwischen Modell und realem System ist dann gegeben, wenn beide eine Struktur aufweisen, wobei jedem Element der Modellstruktur ein und nur ein Element der Realstruktur zugeordnet werden kann. Diese eindeutige Zuordnung muß umkehrbar sein, man nennt sie dann eineindeutig. Vgl. Schröder, K., Hrsg., Mathematik für die Praxis - Ein Handbuch, Bd. I., Frankfurt - Zürich 1964, S. 19. Anderer Auffassung ist Kosiol, E., Modellanalyse als Grundlage unternehmerischer Entscheidungen, in: Zeitschrift für handelswissenschaftliche Forschung, 13. Jg., 1961, S. 318ff, hier S. 321.

mente, die bei der Abbildung des..[Realsystems] durch das Modell gar nicht berücksichtigt werden. Es handelt sich bei der Abbildung des Originals durch das Modell also im allgemeinen nicht um eine Abbildung von M auf M' , sondern um eine Abbildung aus M in M' ." ¹ So enthält z.B. das Modell des komplexen Kaufentscheidungsprozesses als Netzwerk zirkularer Beziehungen (Nicosia-Modell) ² Elemente, die für den Kaufentscheidungsprozeß des Konsumenten irrelevant sind, wie etwaige Spezifizierungen, in welcher Situation das Produkt gekauft wurde. Darüber hinaus enthält das kognitive System "Konsument" Elemente, für die es im Nicosia-Modell kein Urbild gibt, wie etwa den Namen, den Wohnort und die Einkommensklasse des betreffenden Konsumenten. Auch werden gewisse Relationen des Modells und des Konsumenten (des Originals) bei der Abbildung vernachlässigt. Bevor Modell und Realität miteinander verglichen werden können, vereinfacht man beide durch einen Prozeß isolierender Abstraktion. Das Realsystem wird durch ein vereinfachtes System ersetzt, das nur eine Teilmenge von Elementen und Relationen des Realsystems enthält. Erst zwischen dem vereinfachten Realsystem und dessen Abbildung ist es sinnvoll, von Isomorphie und Homomorphie zu sprechen. ³ Während die Isomorphie ein g e t r a u e s Abbild des Realsystems liefert, was bei Modellen niemals der Fall sein wird, bezieht sich die Homomorphie nur auf ein v e r e i n f a c h t e s Abbild. ⁴

1 Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 274.

2 Vgl. Nicosia, F.M., Consumer Decision Processes, Englewood Cliffs, N.J. 1966 und Schulz, R., Kaufentscheidungsprozesse des Konsumenten, a.a.O., S. 61ff. Der komplexe Kaufentscheidungsprozeß als Netzwerk zirkularer Beziehungen (Nicosia-Modell) will u.a. die kaufentscheidungsrelevanten Variablen innerhalb eines Netzes von zirkulären Beziehungen charakterisieren.

3 Vgl. dazu Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S.273.

4 Vgl. ebenda, S. 274.

Ein Modell ist unter dem Gesichtspunkt des theoretisch-inter-subjektiven Ansatzes, der die Probleme der Modellisomorphie und -homomorphie zum zentralen Postulat erhebt, um so eher für die Aufnahme in die Modellbank geeignet, je mehr es die formalisierbaren und intersubjektiv nachprüfbar-
en Abbildungsgegebenheiten erfüllt. Diese Aussage gilt ausschließlich unter dem Aspekt des Verhältnisses zwischen Realsystem (Originalbereich) und Modell (Bildbereich). Das Verhältnis zwischen Benutzer und Modell wird hier nicht beachtet.

Darüber hinaus ist ein Modell um so eher für die Aufnahme in die Modellbank eines Marketing-Informationssystems geeignet, je mehr es die Bedingung der Quantifizierbarkeit von Marketing-Phänomenen erfüllt. Diese Quantifizierbarkeit wird auch u.a. die mengentheoretische Überprüfung seiner Iso- bzw. Homomorphie erleichtern. Im folgenden soll unter Quantifizierbarkeit die Möglichkeit verstanden werden, die für das Modell notwendigen Daten zu beschaffen ("Modellfitting"), um es anhand empirischer Situationen überprüfen zu können. Je mehr sich der Homomorphiegrad in Richtung auf die starke oder die strenge Homomorphie bewegt, um so komplexer wird das Modell. Diese zunehmende Modellkomplexität führt dazu, daß der Modellfehler - d.h. die schwache Homomorphie - abnimmt. Der Zusammenhang zwischen Modellfehler¹ und dem Modellkomplexitätsgrad ist in der folgenden Abbildung idealtypisch als fallende Funktion des Modellfehlers eingezeichnet.

¹ Zum Modellfehler vgl. Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 94 und Ackoff, R.L., Scientific Method - optimizing applied research decisions, New York - London 1962, S. 393ff.

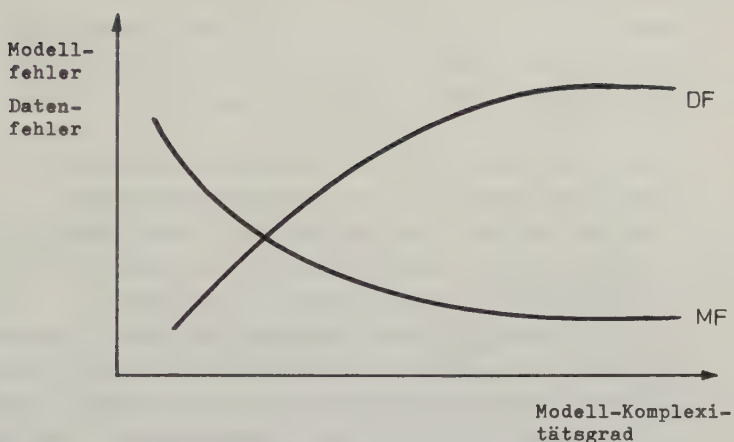


Abb. 16: Modell- und Datenfehler

Geht die Modellkomplexität gegen Null, geht der Modellfehler gegen unendlich, weil das Modell nicht in der Lage ist, das Realsystem hinreichend homomorph abzubilden. Unter dem Gesichtspunkt der Modellhomomorphie muß ein Modell hohen Komplexitätsgrades gewählt werden, das sich dann durch einen geringen Modellfehler auszeichnet.

Komplexe Modelle zeichnen sich aber dadurch aus, daß die Datenbeschaffung als zentraler Gesichtspunkt für die Überprüfbarkeit des Modells problematisch ist. Je komplexer das Modell ist, um so mehr werden Daten benötigt, die a-priori unternehmensintern nicht verfügbar sind und erst durch Beschaffungsaktivitäten zur Verfügung gestellt werden können. Damit ist notwendigerweise ein Datenfehler¹

¹ Vgl. zum Begriff des Datenfehlers Ackoff, R.L., Scientific Method, a.a.O., S. 206ff. Ackoff unterscheidet zwischen Datenfehlern, die auf Beobachtungen zurückgehen, die aus dem Beobachtungsinstrumentarium resultieren, die auf Veränderungen der Umwelt zurückzuführen sind und die sich aufgrund des Beobachtungsfeldes ergeben.

verbunden, der darauf zurückzuführen ist, daß auf originäre und derivative Datenquellen mit fehlerhaften Daten zurückgegriffen werden muß oder gewisse Daten nur innerhalb einer Bandbreite geschätzt werden können.

Von diesem Datenfehler kann angenommen werden, daß er eine steigende Funktion des Modellkomplexitätsgrades ist. Dieser Datenfehler läßt sich in einen Zufallsfehler¹ und in einen systematischen Fehler² unterteilen.

Beim Zufallsfehler, der statistisch darauf zurückzuführen ist, daß nur ein Teil der Daten aus dem Datenkosmos beim "Fitten" des Modells Berücksichtigung findet - die ausgewählten Daten aber fehlerhaft sein können - ist es möglich, daß sich im Zeitablauf die Zufallsfehler ausgleichen.³ Darüber hinaus lassen sich ggf. entsprechend der statistischen Fragestellung Konfidenzintervalle berechnen.

Der systematische Fehler beruht darauf, daß im allgemeinen die ausgewählten Daten eine andere Verteilung der relevanten Merkmalsausprägungen haben als die Daten, die fälschlicherweise erhoben wurden. Über die Größe des systematischen Fehlers lassen sich nur Vermutungen anstellen, und es ist das Grundproblem jeglicher Datenauswahl und -aufbereitung so auszuwählen und aufzubereiten, daß dieser Fehlertyp nach Möglichkeit ausgeschlossen ist. Der systematische Fehler tritt bei der Erhebung, der Aufbereitung (Verschlüsselung, Anfertigung der Datenträger) und bei der Auswertung ein. Er ist untersuchungsobjektabhängig und oft nur durch Erfahrung eliminierbar.

Von dem zur Verfügung stehenden mathematischen Instrumentarium aus ist der Zufallsfehler interessanter, wobei hier

1 Vgl. dazu Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 104ff.

2 Vgl. Deming, W.E., Some Theory of Sampling, New York 1950, S. 24-52.

3 Vgl. Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 52.

zwischen absolutem Fehler¹ (er wird in absoluter Größe vorgegeben) und dem relativen Fehler² (er ergibt sich als Prozentsatz des wahren Wertes) zu unterscheiden ist. Statistisch ist die Auffassung vertretbar, daß es weniger auf den relativen als auf den absoluten Fehler ankommt.³ Werden die zu erhebenden Daten als Stichprobe aus einer unbekannten Grundgesamtheit interpretiert, so führt dies zu dem Schluß, daß die relative Genauigkeit bei den häufig auftretenden Daten wesentlich größer ist als bei den seltenen. Verlangt man für bestimmte Daten einen sehr geringen relativen Fehler, so bedeutet dies, daß der Umfang der zu erhebenden Daten (hier als Stichprobe interpretiert) sehr stark erhöht werden muß, u.U. so stark, daß sich eine Erhebung auf Stichprobenbasis nicht mehr lohnt.⁴

Wenn die Daten so gewählt werden, daß die Anwendung eines geschichteten Stichprobenverfahrens möglich erscheint, lassen sich die einzelnen Stichprobenumfänge theoretisch so festlegen, daß jede Schicht die gleiche Güte, d.h. den gleichen relativen Fehler besitzt.⁵ Darüber hinaus läßt aber die Existenz des systematischen Fehlers es geraten erscheinen, die Berechnungsmöglichkeiten des Zufallsfehlers nicht allzu optimistisch zu beurteilen. Die Forderung, daß jedes Datum die gleiche Chance haben muß, in die Auswahl zu gelangen, führt wohl zur Berechnung des Zufallsfehlers, nicht aber zur Eliminierung des ihn möglicherweise überkompensierenden systematischen Fehlers.

1 Vgl. dazu das Beispiel bei Kellerer, H., Theorie und Technik des Stichprobenverfahrens - Eine Einführung unter besonderer Berücksichtigung der Anwendung auf soziale und wirtschaftliche Massenerscheinungen, Einzelschriften der Deutschen Statistischen Gesellschaft, Nr. 5, 3. Aufl., München 1963, S. 63.

2 Vgl. ebenda, S. 66f.

3 Vgl. Cattell, R.B., The Data Box: Its Ordering Of Total Resources In Term Of Possible Relational Systems, in: Cattell, R.B., Hrg., Handbook of Multivariate Experimental Psychology, Chicago 1966, S. 67ff, hier S. 112.

4 Vgl. zum Stichprobenumfang Kellerer, H., Theorie und Technik des Stichprobenverfahrens, a.a.O., S. 29 und S. 79.

5 Vgl. Deming, W.E., Some Theory of Sampling, a.a.Q, S.230ff.

Durch vertikale Addition des Modell- und des Datenfehlers gelangt man zum Gesamtfehler des Modells. Es wird hier von der Annahme ausgegangen, daß sich der Gesamtfehler additiv aus dem Modell- und dem Datenfehler zusammensetzt. Die Kurve des Gesamtfehlers weist unter Geltung der als idealtypisch unterstellten Kurvenverläufe eine hyperbolische Gestalt auf.

Da der Gesamtfehler $F = F_M + F_D$ ist, beträgt der Durchschnittsfehler

$$f = \frac{F_M + F_D}{K} ,$$

wobei F_M den Modellfehler, F_D den Datenfehler und K den jeweiligen Komplexitätsgrad des Modells bezeichnet. Wird diese Durchschnittsfehler-Funktion nach K differenziert und zur Optimumbestimmung gleich Null gesetzt, so ergibt sich:¹

$$f' = \left(\frac{F_M}{K}\right)' + \left(\frac{F_D}{K}\right)' = 0 .$$

Hieraus folgt, daß

$$\left(\frac{F_M}{K}\right)' = - \left(\frac{F_D}{K}\right)'$$

ist. Das Minimum des Gesamtfehlers liegt somit dort, wo die gespiegelte Kurve des Datenfehlers zur Tangente an die Kurve des Modellfehlers wird. Unter Annahme einer monoton steigenden Kurve des Datenfehlers und einer monoton fallenden Kurve des Modellfehlers liegt das Minimum des Gesamtfehlers bei F_{Min} .

¹ Diese Vorgehensweise ist analog zur Bestimmung der optimalen Betriebsgröße bei Gutenberg. Vgl. dazu Gutenberg, E., Grundlagen der Betriebswirtschaftslehre, Erster Band - Die Produktion, 11. Aufl., Berlin - Heidelberg - New York 1965, S. 427.

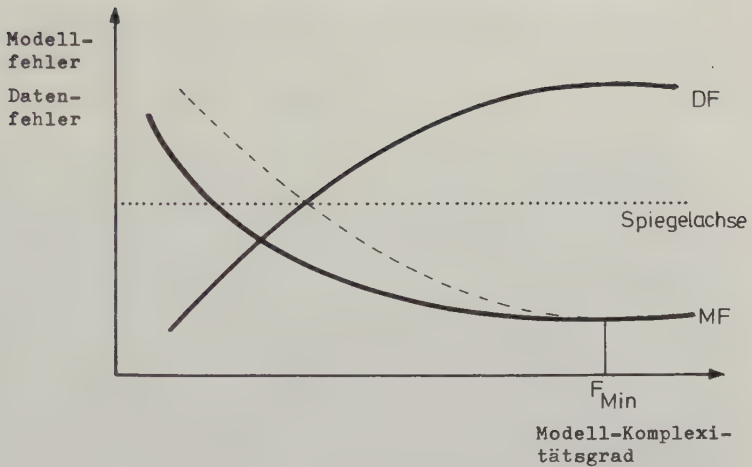


Abb. 17: Bestimmung des Minimums des Gesamtfehlers

Im Gegensatz zur Isomorphie, bei der eine völlige Übereinstimmung zwischen Realität und Modell besteht, liegt bei der Homomorphie nur eine "annähernde" Übereinstimmung vor. Die Modelle, die dem Postulat der Homomorphie genügen, bilden somit nur approximativ die Realität ab. Im allgemeinen sind sie "einfacher" als die Systeme, die sie abbilden, so daß sehr große Komplexität bei den Modellen vermieden wird. Das Postulat der Homomorphie umfaßt zwar sämtliche Anforderungen, die aus theoretischer Warte an Marketing-Modelle gestellt werden können. Da aber die benutzerorientierten Anforderungen hier nicht berücksichtigt werden und das Postulat der Homomorphie nur bedingt geeignet ist, sind weitere Anforderungskriterien zu untersuchen, die einem Entscheidungsträger darüber Auskunft geben, ob ein Modell für die Aufnahme in die Modellbank eines Marketing-Informationssystems geeignet ist oder nicht. In jüngster Zeit ist unter dem benutzerorientierten Postulat ein neuer Ansatz entwickelt worden, der als Gegenthese zum theoretisch-intersubjektiven Ansatz gesehen werden kann.

2.2 Der benutzerorientierte-subjektive Ansatz

Als Gegenthese zum theoretisch-intersubjektiven Ansatz lassen sich die von LITTLE¹ zusammengefaßten und zum "Decision Calculus" vereinigten Gedankengänge aufstellen.

Die Ausgangsüberlegung von LITTLE basiert auf der Tatsache, daß das große Dilemma von Marketing-Modellen darin besteht, daß sie vom Marketing-Management nicht angewendet werden.² Dies hat u.a. seinen Grund darin, daß die Marketing-Entscheidungssträger nicht in unmittelbarem Kontakt mit dem Modell stehen, es nicht entwickelt haben, seine Arbeitsweise nicht begreifen und folglich bei ihnen das Zutrauen und die Identifizierung mit diesem Modell nur rudimentär oder gar nicht vorhanden sind. Zum gegenwärtigen Zeitpunkt ist bei der Entwicklung von Marketing-Modellen zu erwarten, daß wir "von solchen Modellen am meisten für die Entscheidungsverbesserung erwarten dürfen, die dem Entscheidungssträger ermöglichen, die Gesamtheit der ihm zur Verfügung stehenden Daten, Meinungen und Erwartungen so zu integrieren, daß das Modell als eine Erweiterung des unternehmerischen Denkprozesses verstanden werden soll."³ In diesem Sinn ist ein Decision-Calculus-Modell ein Programm zur Verarbeitung der einem Entscheidungssträger zugehenden Informationen, das durch ein charakteristisches Muster der Informationsverarbeitung geprägt ist, wobei dieses Muster durch Erfahrung, Vorschrift, Experiment oder Tradition hergebracht wurde.

Bevor die Gestaltungsprinzipien von Decision-Calculus-Modellen erörtert werden sollen, ist es zweckmäßig, den Hintergrund zu untersuchen, warum der Ansatz von LITTLE als

1 Vgl. dazu Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, in: Management Science, Vol. 16, 1970, S. B-466ff.

2 Vgl. ebenda, S. B-466.

3 Krautter, J., Marketing-Entscheidungsmodelle, a.a.O., S. 63.

ein neuer Ansatz zur Modellselektion und -entwicklung und nicht als eine Aufzählung mehr oder weniger bekannter Gestaltungsprinzipien anzusehen ist.

Bislang ist den Managementfunktionen im Marketing der Prozeß der Selektion und Entwicklung von Marketing-Modellen konträr entgegengesetzt. Er ist nicht als Routineprogramm, sondern als "Zweckprogramm"¹ anzusehen, für dessen Implementierung zwar die Ziele vom Management vorgegeben sind, zu dessen Verwirklichung aber jedoch erst die Modelle und ihre zweckentsprechenden Einsatzmöglichkeiten gefunden werden müssen. Der Ansatz von LITTLE muß vor dem Hintergrund gesehen werden, daß die sich dem Marketing-Management stellenden Probleme als "interaktive Modelle" abzubilden sind, die einerseits durch eine Struktur der miteinander kommunizierenden, informationsverarbeitenden Rollen, andererseits durch die enge zeitliche Überlagerung von vier logischen Planungsphasen geprägt sind. Jede dieser Planungsphasen wird durch einen spezifischen Zweck der Informationsverwendung und -verarbeitung charakterisiert.

Die Aufgabe der Modellselektion und -entwicklung im Bereich des Marketing läßt sich als ein Zweckprogramm besonderer Ausprägung interpretieren. Ein Modell entsteht aus einem marketingimmanenten Problem heraus, wodurch die Zwecksetzung des Modells nun sehr umrißhaft festgelegt wird. Die anfänglichen Zwecke sind noch nicht derart konkretisiert, daß sie als konsistentes Zielsystem formuliert werden können.² Darüber hinaus lassen sich aus diesen anfänglichen Zwecken noch keine konkreten Anhaltspunkte über den Modelltyp und das anzuwendende mathematische Verfahren festlegen. Es ist vielmehr Aufgabe eines derartigen Zweckprogramms, die Ziele insoweit zu konkretisieren, daß ein

¹ Vgl. dazu Luhmann, N., Zweckbegriff und Systemrationalität - Über die Funktion von Zwecken in sozialen Systemen, Tübingen 1968, S. 326.

² Vgl. ebenda, S. 215ff.

Modell innerhalb eines akzeptablen und realisierbaren Lösungsbereichs selektiert bzw. entwickelt werden kann. Daraus folgt, daß erst in weiteren Lösungsstadien von den Modellkonstrukteuren Entscheidungen über die mit dem Modell verfolgten Absichten gefällt werden. Bis dahin ist der Modellzweck oft latent, niemand spricht über ihn und bei seiner Aufdeckung könnten schwerwiegende Konflikte und Differenzen über den Sinn dieses Modells entstehen. Die schließlich verabschiedete Richtung der Modellentwicklung ist schließlich auch nur wiederum ein Zwischenschritt, der für die weitere Vorgehensweise Prämissen setzt und ist für Problemlösungsverhalten in komplexen Situationen typisch.¹ Das folgende Beispiel möge dieses näher verdeutlichen.

Das Marketing-Management steht vor dem Problem, ein neues Produkt am Markt einzuführen und mittels eines Modells die Auswirkungen auf Umsatz, Marktanteil und Gewinn zu prognostizieren. Dieser Zweck ist vorgegeben, allgemein akzeptiert und einem Produktmanager zur Lösung übertragen. Welche Zwecke nun im einzelnen bei dieser Problemlösung verfolgt werden, ist weder einheitlich feststellbar noch konsistent nachzuvollziehen. Aus dem Kreis der Produktmanager werden Zwecke auf unterschiedlichen Abstraktionsebenen vielfach widersprüchlich und ohne inneren logischen Zusammenhang vorgetragen. Das Problem, vor dem der mit der Lösung beauftragte Produktmanager steht, ist vieldimensional und komplex.

An dieser Stelle scheint es angebracht zu sein, den Begriff "Problem" näher zu präzisieren. "Ein Problem ist gegeben, wenn ein Individuum ein bestimmtes Ziel erreichen will, jedoch nicht weiß, wie es zu diesem Ziel gelangen kann, also nicht auf spezifische Verfahren, Techniken und Operationen zurückgreifen kann."² Problemsituationen, die mittels Mo-

1 Vgl. Luhmann, Politische Planung, in: Jahrbuch für Sozialwissenschaften, 17.Jg., Bd. 1, 1966, S. 271ff, hier S. 272.

2 Süsswold, F., Bedingungen und Gesetzmäßigkeiten des Problemlösungsverhaltens, in: Graumann, C.E., Hrsg., Denken, 4. Aufl., Berlin 1969, S. 273ff, hier S. 273.

dellen einer Lösung nähergebracht werden sollen, zeichnen sich vor allem dadurch aus, daß das zu entwickelnde Modell nur als vage Vorstellung bei den Modellkonstruktёuren und beim Management vorhanden ist.

Für die weitere Untersuchung soll festgehalten werden, daß ein Modellselektions- und -entwicklungsprozeß dahingehend interpretiert werden kann, daß einerseits Komplexität reduziert, andererseits mit zunehmender modellmäßiger Strukturierung des Marketingproblems Komplexität aufgebaut wird.¹ Die Komplexitätsreduktion ergibt sich beim angeführten Beispiel daraus, daß es den Modellkonstruktёuren gelingt, die empirischen Probleme zu erfassen, während gleichzeitig das sich in Entwicklung befindliche Modell an Komplexität gewinnt.

Dieser hier nur umrißhaft skizzierte Prozeß kann in vier logische Phasen aufgelöst werden, die sich durch die Verarbeitung der den Entscheidungsträgern zur Verfügung stehenden Daten, durch die Kommunikation mit dem Marketing-Management und durch die Mitwirkung des Managements an der Modellentwicklung voneinander abgrenzen lassen.²

1. In der Phase der Problemstrukturierung geht es darum, Erkenntnisse und Einsichten über das zu lösende Problem zu gewinnen. Die Verarbeitung vorhandener Daten und die Mitwirkung des Marketing-Management dienen primär diesem Zweck.
2. In der Phase der Lösungssuche wird das Modell von den Modellkonstruktёuren entwickelt. Modellspezifische Informationen dienen dazu, die entwickelten Lösungsvorschläge kritisch zu beurteilen.

1 Vgl. Luhmann, N., Zweckbegriff und Systemrationalität, a.a.O., S. 215.

2 Zu diesen Phasen vgl. etwa Naylor, T.H., Marketing-Politik-Modelle, o.O. und o.J., S. 8ff.

3. In der Phase der Lösungsanpreisung muß das Marketing-Management von den Vorteilen, der Praktikabilität etc. des Modells überzeugt werden. In dieser Phase dienen Informationen primär dazu, das Management mit den Vorteilen des Modells bekanntzumachen, um es "verkaufen" zu können.
4. Schließlich erfolgen in der Phase der Lösungskonkretisierung Weisungen des Marketing-Managements an die Modellkonstrukteure über die weitere Entwicklungsrichtung des Modells, und zwar aufgrund der sich bislang eingestellten Modellergebnisse. Information wird in dieser Phase überwiegend zu Kontrollzwecken verwendet.

Für die Abfolge dieser Planungsphasen wird unterstellt, daß sie in enger Übereinstimmung mit dem Prozeß der Modellselektion und -entwicklung stehen und damit für das interaktive Planungsmodell konstituierend wirken.

In der Phase der Problemstrukturierung wird z.B. ein Marketingproblem vom Management erkannt. Die Problemstrukturierung verläuft dabei schrittweise, wobei Beobachtungen, Daten und andere Informationen durch Systematisierung, Ordnung und Begriffsbildung gedanklich verarbeitet werden. Durch derartige Informationsverarbeitungsaktivitäten wird ein begriffliches Modell, das als "internes Umweltmodell"¹ bezeichnet werden soll, bei jedem in dieser Phase Beteiligten aufgebaut. In dem Maß, wie die Komplexität des als problematisch angesehenen Phänomens durch Problemstrukturierung abgebaut wird, gewinnt das interne Umweltmodell an Komplexität. Die Gewinnung von Informationen, verstanden als Aussage über die Existenz, die Struktur und das Verhalten realer Phänomene, bei denen eine Überprüfung möglich ist, erfolgt stets mittels kognitiver Modelle.² Die wissen-

1 Klaus, G., Hrsg., Wörterbuch der Kybernetik, 2. Aufl., Berlin 1968, S. 287.

2 Vgl. Kirsch, W., Entscheidungsprozesse, Bd. II, a.a.O., S. 138ff.

schaftliche Form des internen Umweltmodells entspricht den "erfahrungswissenschaftlichen Theorien"¹, während das mehr vorwissenschaftliche empirische Ordnungssystem eine mehr oder weniger straff geordnete Vielfalt von Einzelinformationen enthält, die nach "erlebter Kausalität"² geordnet wird. Neue Informationen können aufgrund des internen Umweltmodells nur dann gewonnen werden, wenn sie ihm widersprechen bzw. sich nicht in das interne Ordnungssystem einordnen lassen. Was dann als potentiell neue Information zu gelten hat, kann demnach nur in bezug auf ein bestimmtes internes Ordnungssystem entschieden werden.

Die Beseitigung von Widersprüchen zwischen internem und externem Modell ist nur über eine Umstrukturierung des internen Ordnungssystems zu erreichen, d.h. es wird eine Anpassung des internen Modells an die realen Phänomene vorgenommen. So hat z.B. der Marketing-Manager einen Erfahrungsschatz, der ihn befähigt, bestimmte Marketing-Phänomene einzuordnen. Tritt nun ein Ereignis am Markt ein, über das er keine Erfahrung besitzt, wird es ihm nur dann gelingen, dieses neue Ereignis einzuordnen, wenn er bereit ist, sein internes Umweltmodell (seinen Erfahrungskomplex) zu revidieren.

Der Aufbau von internen Umweltmodellen kann als ein Prozeß verstanden werden, der zur Entwicklung von Marketing-Modellen parallel verläuft. Jegliche Form rationaler Informationsgewinnung und -verarbeitung baut auf der skizzierten Vorgehensweise auf. Sie soll als erstes Postulat rationaler Informationsgewinnung und -verarbeitung bezeichnet werden. Dieses Postulat besagt, daß gültige Informationen nur dann gewonnen werden können, wenn die laufende Kontrolle der informationsgewinnenden kognitiven Aktivitäten durch die Beobachtung der Realität gewährleistet ist.³ Die Not-

1 Albert, H., Probleme der Theoriebildung, in: Albert, H., Hrsg., Theorie und Realität, a.a.O., S. 27.

2 Stachowiak, H., Denken und Erkennen im kybernetischen Modell, 2. Aufl., Wien-New York 1969, S. 142.

3 Vgl. Albert, H., Probleme und Theoriebildung, in: Albert, H., Hrsg., Theorie und Realität, a.a.O., S. 22ff.

wendigkeit der Beobachtung der Realität ergibt sich insbesondere bei der Selektion von Marketing-Modellen, da jeder Entscheidungsträger tendenziell geneigt ist, vorwissen-schaftliche Urteile aufgrund seines internen Umweltmodells zu fällen. Somit ist es originäre Aufgabe der Marketing-Theorie, die "Alltagserfahrung" und die "Interessenideologie" der Entscheidungsträger durch Analyse der Probleme zu überwinden. Dabei steht es keineswegs so, daß die inter-essenbedingte Alltagserfahrung unrichtig sein muß. Sie kann recht und sie kann unrecht haben. Um dieses zu entscheiden, ist ein strenges Kriterium und eine wissenschaftliche Methode zu finden.¹

Letztlich beruht dieses Postulat auf der Tatsache, daß die mittels kognitiver Aktionen gewonnenen Aussagen über die Realität hypothetischen Charakter besitzen, so daß die Gültigkeit dieser Hypothesen anhand der Tatsachen selbst geprüft werden muß. Das erste Postulat rationaler Informationsgewinnung läßt sich somit als Rückkoppelungspro-zeß zwischen kognitiv und real beobachtbaren Phänomenen interpretieren.

Darüber hinaus soll die Prüfung der Ergebnisse der kognitiven Aktionen gewährleisten, daß sich das durch diese Aktionen gewonnene "Neue" auch bewährt hat. Durch den Vergleich der durch kognitive Aktionen gewonnenen hypothetischen Aussagen mit der Realität ist es möglich festzustellen, ob die Aussage etwas Neues, d.h. etwas das interne Umweltmodell Umstrukturierendes enthält, das dann entweder als "potentiell neue Information"² zu gelten hat oder nicht. Führt die Aussage zu einer Umstrukturierung des internen Umweltmodells eines Marketing-Managers, ist sie wert, aufbewahrt zu werden, bis neue hypothetische Aussagen umfassenderer Art sie ablösen. In diesem Fall erweisen sich die neuen hypothetischen Aussagen als genereller und damit leicht-

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 4.

2 Hofstätter, P., Psychologie, Frankfurt 1957, S. 168.

ter falsifizierbar, so daß sie unter dem Gesichtspunkt der empirischen Prüfbarkeit den ersteren hypothetischen Aussagen überlegen sind. Dies führt zu dem Schluß, daß Marketing-Manager mit einem generellen internen Umweltmodell besser in der Lage sein werden, neue Informationen zu verarbeiten, als solche, deren internes Umweltmodell auf speziellen Gegebenheiten basiert.

Das zweite Postulat rationeller Informationsgewinnung und -verarbeitung ergibt sich aus der Tatsache, daß jeder Satz bzw. jedes System von Sätzen vorläufigen Charakter hat.¹ Für den Prozeß der Informationsgewinnung und -verarbeitung folgt hieraus, daß bei Kongruenz zwischen Explikans und Explikandum der Informationsgewinnungs- und -verarbeitungsprozeß nicht als abgeschlossen gesehen werden darf. Der Prozeß der Falsifizierung ist solange fortzusetzen, bis ein empirisches Phänomen in Widerspruch zur betreffenden Ausgangshypothese steht, so daß der Prozeß der Hypothesenprüfung von neuem beginnen kann.

Demnach ist für die Vermehrung der bereits vorhandenen und geprüften Informationen nicht das Auffinden hypothesenkonformer, sondern das Finden von hypothesenkonträren Situationen von Interesse.² Angesichts der Tatsache, daß das Marketing-System ein offenes sozio-technisches System ist, das sich in ständigen Transaktions- und Interaktionsbeziehungen mit seiner Umwelt befindet, führt die Änderung der Umweltbedingungen schon allein dazu, daß sich potentiell neue Informationen aufgrund der Variation der Ausgangsbedingungen ergeben.

Aus diesem Gesichtspunkt heraus ergibt sich das zweite Postulat rationeller Informationsgewinnung: Die Vergrößerung

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 225.

2 Vgl. derselbe, Truth, Rationality, and the Growth of Scientific Knowledge, in: Popper, K.R., Hrsg., Conjectures and Refutations - The Growth of Scientific Knowledge, New York - London 1962, S. 242ff, hier S. 242.

des Bestandes an bewährten Informationen¹ erfordert die Prüfung der Hypothesen unter den jeweiligen Bedingungen, denen sich der Informationsempfänger zum Zeitpunkt der Prüfung gegenübersteht. Dabei sind jene Umweltkonstellationen für das Marketing von besonderem Interesse, die extrem gelagert sind (z.B. Währungskrisen, staatliche Eingriffe in die Wirtschaftspolitik), die sich also durch eine geringe Prüfbarkeit auszeichnen.

Das sich in der Phase der Problemstrukturierung langsam aufbauende interne Umweltmodell kann als eine begriffliche Gesamtheit verstanden werden, die aus einer Anzahl mehr oder weniger integrierter, auf einzelne Aspekte eines Problems bezogenen Konzepten gebildet wird. Da sich jedes ursprünglich als "einfach" verstandene Marketing-Problem als Syndrom erweisen kann, bei dem "alles mit allem zusammenhängt", sind bei der Problemstrukturierung Grenzen anzulegen. Dadurch werden bestimmte Aspekte des Marketing-Problems aus der Betrachtung ausgeschlossen. Darüber hinaus sind Annahmen einzuführen, die einer einfacheren Handhabung des Problems dienen sollen.

Zwischen den an der Entwicklung von Marketing-Modellen Beteiligten muß eine Konsensbildung darüber stattfinden, daß die jeweils möglichen Problemlösungskonzepte der Behandlung von Marketing-Problemen angemessen sind. Kontinuierliche Kommunikation zwischen Marketing-Management und den Modellkonstrukteuren und eine Angleichung an einen gemeinsamen Informationsstand wird unerlässlich. Konzepte aus dem internen Umweltmodell werden dabei durch Aussagen zwischen Management und Modellkonstrukteuren kommuniziert, die je nach den kommunikativen Notwendigkeiten auf unterschiedlichen Abstraktionsebenen vorgenommen werden. So kann z.B. eine Aussage, daß 6 von 10 Rauchern den Tabak XY kaufen werden, in die allgemeinere Aussage, daß der Marktanteil vermutlich hoch

¹ Vgl. Lenk, H., Läßt sich die Erkenntnis kybernetisch begründen?, in: Göttingische gelehrte Anzeigen, 219. Jg., 1967, S.131ff, hier S. 146ff.

sein wird, überführt werden. Diese hohe Auslösbarkeit von Begriffen in einer Kommunikationssituation, zusammen mit der schnellen begrifflichen Zusammenführung divergierender Aspekte, zeichnet interne Umweltmodelle vor Modellen in mathematischer Formulierung aus.¹ Sollen mathematische Modelle eine Kommunikation zwischen Modellkonstrukteuren und Marketing-Managern stützen, so bedürfen sie erst einer Rückübersetzung in solche Marketing-Aussagen, die der Struktur der internen Umweltmodelle der Marketing-Manager angemessen sind.

Die Strukturierung des internen Umweltmodells ist darüber hinaus von der Marketingerfahrung, der fachlichen Ausbildung und der Verständnisbereitschaft der an der Modellselektion und -entwicklung Beteiligten abhängig. Die Schlussfolgerungen in der Phase der Problemstrukturierung werden um so detaillierter sein, je stärker diese das interne Umweltmodell prägenden Faktoren sein werden.

In der sich anschließenden Phase der Lösungssuche erarbeiten die Modellkonstrukteure unter kommunikativer Mitwirkung des Marketing-Management Problemlösungen. Diese Phase soll als Komplement zur Phase der Problemstrukturierung verstanden werden. Sie ist Teil eines projektgebundenen Lernprozesses. Je mehr sich die Modellkonstrukteure an eine realisierbare und akzeptable Lösung "herantasten", um so besser lernen sie das Problem kennen, entdecken neue Aspekte und Grenzen und sehen, welche Lösungsrichtung keinen Erfolg verspricht. Alle diese akzeptablen und realisierbaren Lösungen liegen innerhalb eines begrenzten Suchraumes, der teilweise vorher bekannt ist, teils aber erst mit zunehmender Lösungskonkretisierung - mit dem Aufdecken der Ziele und Randbedingungen - quasi als Teil eines projektgebundenen Lernprozesses deutlich wird. Die kommunikative Mitwirkung des Marketing-Management kann diesen Lernprozeß erheblich beschleunigen.

¹ Vgl. Naschold, P., Systemsteuerung, Stuttgart 1969, S. 78ff.

Die Lösungssuche ist iterativ. Sie beginnt mit einer detaillierten Skizzierung und Festlegung von Annahmen, die bei fortschreitendem Lösungsprozeß verdichtet und konkretisiert werden. Führt ein einmal eingeschlagener Weg zu keiner akzeptablen oder realisierbaren Lösung, dann wird zu den ursprünglichen Annahmen zurückgekehrt. Sie werden modifiziert und in einer anderen Richtung präzisiert.

Die Länge der Suche nach einer oder mehreren akzeptablen und realisierbaren Lösungen hängt von der Komplexität des zu lösenden Problems, von der Breite des Suchraumes, von der bereits erworbenen Routine und von der politischen Lage im Unternehmen ab. Die Lösungssuche wird in immer neuen Iterationen so lange weitergeführt, bis sich eine akzeptable und realisierbare Lösung ergibt. Dieser Prozeß ist in der folgenden Abbildung dargestellt.

Ausgangspunkt dieser Abbildung ist ein zu lösendes, d.h. in einem Modell abzubildendes Marketingproblem, wobei nach den Zwischenschritten der Abstraktion, der Bildung eines internen Umweltmodells und der Klassifizierung des Problems eine Analyse des zu lösenden Problems erfolgt. Das Problem wird anschließend explizit formuliert, um unter Beachtung von beschafften oder zu beschaffenden Daten, Hypothesen über die Lösung aufzustellen und diese dann zu testen. Bei der Datensammlung erfolgt zunächst eine Klassifizierung zwischen lösungsrelevanten und lösungsirrelevanten Variablen des zu entwickelnden Modells, wobei die relevanten Variablen zu Inputinformationen bei der Hypothesenbildung werden. Schließlich muß es gelingen, Verknüpfungen zwischen den Variablen zu entdecken und Unterprobleme zu formulieren, für die analytische oder heuristische¹ Lösungsverfahren zur Verfügung stehen. Die Stufe des Hypothesentests und die Lösung von Unterproblemen stehen in Rückkoppelung zur Problemanalyse.

¹ Bei diesen heuristischen Lösungsverfahren handelt es sich um solche im Asoffschen Sinne.

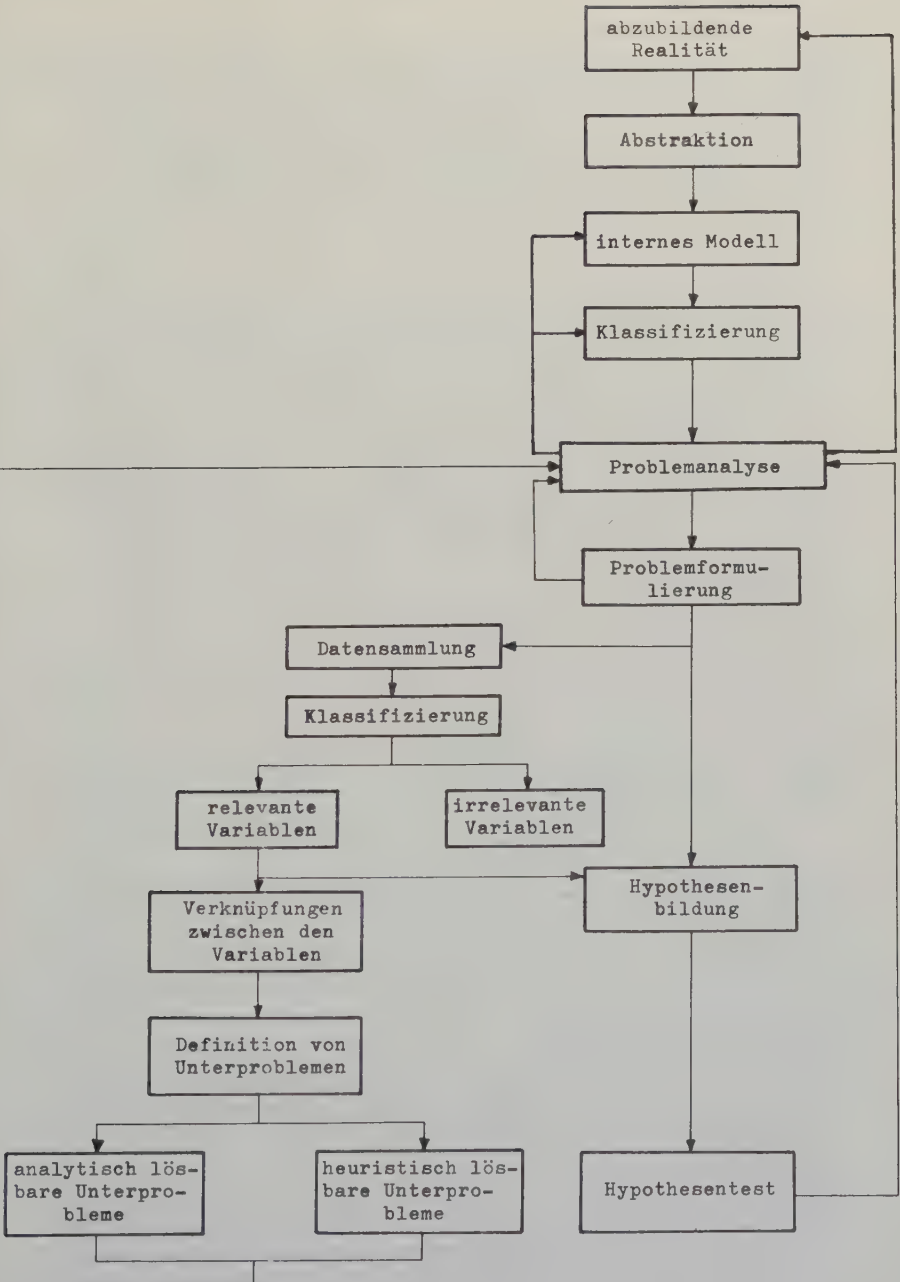


Abb. 18: Problemlösungsprozeß

Unter den verschiedenen Gründen, die zu einer verstärkten Anwendung dieser Vorgehensweise führen können, ist eine der häufigsten eine "Art von Mißbehagen" gegenüber den Kriterien, welche jeweils bestimmen, was als eine "befriedigende Erklärung"¹ eines Marketing-Phänomens zu gelten hat.

In jüngster Zeit wird häufig die Forderung nach der Entwicklung von Marketing-Modellen gestellt, die durch ständige Überprüfung an der Realität verbessert werden können, so daß ihr Wert im Zusammenhang mit den Gestaltungsaufgaben² des Marketing erhöht wird. Der Wert eines Marketing-Modells "kann daher, zumindest soweit .. [es] Entscheidungen von Führungsproblemen absichern.. [soll] im allgemeinen nur subjektiv vom Manager selbst erlassen werden, und zwar in bezug darauf, wie das Modell dem Manager hilft, das Entscheidungsproblem klarer und strukturierter zu sehen, die das Problem beeinflussenden Faktoren in ihrer Wirkung besser überschauen sowie sämtliche verfügbaren Daten verarbeiten zu können."³ Bei Prognosemodellen steht wie auch bei den Entscheidungsmodellen nicht primär der subjektive Wert für einen Marketing-Manager im Vordergrund, sondern die Objektivität und Genauigkeit des Prognoseergebnisses.

In der Phase der Lösungssuche spielt Information zum Zweck der Beurteilung des entwickelten Modells für das Management eine bedeutsame Rolle. Will man, nachdem ein Modell entwickelt worden ist, feststellen, wie es sich in einer kon-

1 Eine Erklärung kann als "befriedigend" angesehen werden, wenn Kriterien bekannt sind, welche die Mindestanforderungen spezifizieren und die in Frage kommende Erklärung diesen Anforderungen genügt. Vgl. dazu die Definition des Zielspruchsniveaus bei Heinen, E., Das Zielsystem der Unternehmung, a.a.O., S. 83.

2 Zu den betriebswirtschaftlichen Gestaltungsaufgaben vgl. Kappler, E., Systementwicklung - Lernprozesse in betriebswirtschaftlichen Organisationen, Bd. 11 der Schriftenreihe "Die Betriebswirtschaftslehre in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1972, S. 121ff.

3 Krautter, J., Marketing-Entscheidungsmodelle, a.a.O., S. 30 (Im Original gesperrt).

kreten Marketingsituation bewährt hat, können operationale Bewertungsverfahren herangezogen werden. Da der Suchprozeß nach Lösungen iterativ ist, werden bei jeder Modifikation eines Lösungsvorschlages die gleichen Verfahren und Ausgangsdaten in immer neuen Zusammenstellungen verwendet.

Sowohl für Prognose- als auch für Entscheidungsmodelle stehen insbesondere die Probleme der Beurteilung dieser Modelle aufgrund der mit ihnen gefundenen Lösungen im Vordergrund. Die Modelle sollen gewährleisten, daß eine sachlogisch adäquate Problemformulierung und -lösung ermöglicht wird. Diesem Postulat kommt die vorhandene Modellpluralität entgegen, die dazu führt, daß qualitativ bessere Entscheidungen und Prognosen aufgrund mehrerer Modelle ermöglicht werden, die dann miteinander verglichen werden können.

Unter Modellpluralität soll die Tatsache verstanden werden, daß zur Lösung eines gegebenen Problems mehrere Modelle existieren, die sich teilweise durch die Komplexität der Modellstruktur, teilweise durch das zur Anwendung gelangende mathematische Lösungsverfahren unterscheiden. Unter Berücksichtigung der bei vielen Modellen zur Verfügung stehenden ontologischen und nomologischen Hypothesen ist eine Entscheidung oder Prognose eher möglich als bei nur wenigen vorhandenen Modellen. Darüber hinaus ermöglicht es der weit gespannte Rahmen möglicher Hypothesen, aufgrund verschiedener Marketing-Modelle Probleme unterschiedlich zu lösen. Stimmen Prognosen und Lösungen nicht mit der empirischen Erfahrung des Managements überein, besteht aufgrund der Modellpluralität durchaus die Möglichkeit, mit anderen Modellen zu besseren Resultaten zu gelangen.

Die Ursachen der Modellpluralität sind erstens in den unterschiedlichen Prämissen zu suchen, die bei der Modellkonzeption zugrunde gelegt werden. Die Voraussetzungen, welche mittels der Prämissen über die Wirklichkeit und

Reichweite von Erkenntnisobjekten gemacht werden, dienen einerseits dazu, den Bedingungsrahmen festzulegen, unter dem eine Erklärung als "vernünftig" gelten kann. So sind Marketing-Modelle, die beispielsweise für eine Einproduktunternehmung entwickelt worden sind, nur für derartige Unternehmen sinnvolle Erklärungsgrundlagen. Schlüsse für Mehrproduktunternehmen lassen sich aus ihnen nicht ableiten, ohne daß gravierende Fehlaussagen in Kauf genommen werden. Darüber hinaus haben die Voraussetzungen die Funktion, die theoretische Marketing-Forschung in eine bestimmte Richtung zu lenken. Den Voraussetzungen kommt in diesem Fall eine strukturierende und ordnende Funktion von Prognose- und Entscheidungsproblemen zu.

Weiterhin haben die Prämissen Konsequenzen bezüglich des zur Anwendung gelangenden mathematischen Instrumentariums (z.B. der Marginalanalyse oder den Verfahren der mathematischen Programmierung) und des in die Analyse einbezogenen Entscheidungsverhaltens der Entscheidungsträger im Marketing. Die Einbeziehung des Entscheidungsverhaltens führt dazu, die Prämisse der unbeschränkten Informationsverarbeitungskapazitäten der Entscheidungsträger zugunsten des weitaus realistischeren beschränkt-rational intendierten Entscheidungsverhaltens aufzugeben.

Der zweite Grund der Modellpluralität liegt mehr auf pragmatischer Ebene. Bedingt durch die zunehmende Dynamik wirtschaftlichen Geschehens, ändern sich die zu lösenden Probleme häufig. Somit bedarf es des Prinzips der isolierenden Abstraktion, um die wesentlichen Determinanten von Marketing-Phänomenen herauszufinden. Daraus ergibt sich, daß angesichts der Fülle zu erklärender Phänomene im Marketing durchaus verschiedene Abbildungen -d.h. Modelle - existieren können.¹

¹ Zu den entsprechenden wissenschaftstheoretischen Gesichtspunkten vgl. Bochenski, I.M., Die zeitgenössischen Denkmethode, 3. Aufl., Bern - München 1965, S. 36ff und S. 51ff.

An die Phase der Lösungssuche schließt sich die Phase der Lösungsanpreisung an. Die in der Phase der Lösungssuche gefundene Lösung muß hier kommuniziert werden. Insbesondere ist Kommunikation mit dem Marketing-Management vorzunehmen, da dieses die Lösung verstehen muß, damit es in ihre Verwirklichung Zutrauen besitzt.¹ Die Kommunikation mit dem Marketing-Management hat insofern eine Rückkopplungsfunktion, als daß erst gegenüber einem Lösungsvorschlag Informationen über die vielfältigen, latenten, oft unkoordinierten Ziele vom Management gegeben werden. Dieses führt zuweilen dazu, daß die Lösung nicht akzeptabel oder in einer bislang nicht berücksichtigten Richtung zu suchen ist.

Die Kommunikation einer Lösung wird daher von den Modellkonstrukteuren nicht ohne Vorbereitung und taktisch geschicktem Vorgehen vorgenommen, da sie danach trachten, ihre Lösung gegenüber den Marketing-Management durchzusetzen. Daraus folgt, daß die Lösung "angepriesen" werden muß. Das Anpreisen von Lösungen erfordert zwei Arten von Informationen. Zunächst werden "Verkaufsinformationen" benötigt, mit deren Hilfe man die Vorzüge einer Lösung in den Vordergrund des Interesses "spielen" und bereits bestehende Lösungen ggf. "disqualifizieren" kann. Darüber hinaus sollen die "Verkaufsinformationen" so beschaffen sein, daß sie insbesondere auf die Realisierbarkeit der Lösung abgestellt sind. Die zweite Art benötigter Information kann mit "Intelligenzinformation"² bezeichnet werden. Sie wird dazu benötigt, um die Einstellung des Marketing-Management bzw. anderer Adressaten der Lösung gegenüber dieser Lösung abschätzen zu können. D.h. mit anderen Worten, daß Informationen über die internen Umweltmodelle der Entscheidungsträger gewonnen werden müssen, die über die Annahme oder Ablehnung der Lösung zu entscheiden haben. Erst wenn mit-

1 Vgl. dazu Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-467.

2 Vgl. Webber, M.M., The Role of Intelligence Systems in Urban Planning, in: Journal of American Institute of Planners, November 1965, S. 289ff, hier S. 289.

tels "Intelligenzinformation" Kenntnisse über die internen Umweltmodelle der Adressaten der Lösung gewonnen wurden, können "Verkaufsinformationen" gezielt eingesetzt werden, damit sich das interne Umweltmodell dieser Marketing-Manager dergestalt ändert, daß sie bereit sind, die neue Lösung zu akzeptieren. Information wird in der Phase der Lösungsanpreisung primär zum Zweck der Manipulation und zum Durchsetzen der Interessen der Modellkonstrukteure verwendet. Problem dabei ist ihre zeitgerechte Bereitstellung und die sorgfältige Überprüfung ihrer Aussagekraft, damit sich bei dem Marketing-Management keine Fehleinschätzungen ergeben und der manipulatorische Zweck erreicht wird. NAYLOR¹ stellt dazu fest, daß die Modellkonstrukteure oft nicht sehr gute "Verkäufer" ihrer eigenen Modelle sind, da sie oft Schwierigkeiten haben, die Ergebnisse ihrer Modelle dem Marketing-Management zu erläutern. In jeder Hinsicht ist die Politik des Durchsetzens von Modellen gegenüber dem Management der problematischste Teil der Modellentwicklung. Es ist in der Tat durchaus möglich, daß ein brauchbares Modell vom Marketing-Management der Unternehmung "abgeschossen" wird, weil der Modellkonstrukteur nicht über die Eigenschaft verfügt, sein Modell anzupreisen.

In der Phase der Lösungskonkretisierung werden Bedingungen für das weitere detaillierte Vorgehen gesetzt.² So dürfte z.B. bei einem Marketing-Mix-Modell kein derart ausgereifter Vorschlag zur Verfügung stehen, der nicht verbesserungsfähig wäre. In dieser Phase treten Schwierigkeiten auf, die primär auf unkontrollierbare Einflüsse exogener Variablen zurückgehen. Es ist somit Information zum Zweck der Kontrolle, zur Beurteilung der Auswirkungen und des Fortschritts bei der Realisierung der Lösung erforderlich. Für derartige Zwecke können Indikatoren als Maßstäbe der Zielerfüllung formuliert und die Abweichungen im Sinn einer regelmäßigen Berichterstattung festge-

¹ Vgl. Naylor, T.H., Marketing-Politik-Modelle, a.a.O., S. 40.

² Vgl. Luhmann, N., Politische Planung, a.a.O., S. 272.

stellt und kommuniziert werden. Es kann dann vom Marketing-Management korrigierend eingegriffen werden, wenn sich erhebliche Abweichungen herausstellen und die Verwirklichung der Lösung in Gefahr gerät. Die Verwendung dieser Information ist ähnlich einem konditionalen Programm: Wenn eine Abweichung vom Standard gemeldet wird, dann sind neue Entscheidungen seitens des Marketing-Management erforderlich. Hier gestatten systematisch vorgenommene Beobachtungen von Indikator-Zeitreihen, Schlüsse auf die Wirkungsweise bestimmter berücksichtigter und nicht berücksichtigter Variablen zu ziehen.

Erst unter Beachtung der Prozesse in den vier beschriebenen Planungsphasen erscheinen die Gestaltungsprinzipien von Decision-Calculus-Modellen sinnvoll und zeigen, daß es sich bei diesen aus den vier Planungsphasen abzuleitenden Kriterien um einen neuen Ansatz zur Modellselektion und -entwicklung handelt, der dem theoretisch-inter-subjektiven Ansatz gegenübersteht.

Die Gestaltungsprinzipien umfassen nach LITTLE¹ folgende:

1. E i n f a c h h e i t : Das Prinzip der Einfachheit hat die Aufgabe, das Verständnis des Modells zu erleichtern. Einfachheit ist dann gegeben, wenn bedeutsame Phänomene im Modell berücksichtigt, unbedeutende vernachlässigt werden. Die Aufnahme von detaillierten Phänomenen sollte so lange im Modell nicht erfolgen, bis der Benutzer bereit ist, diese zu berücksichtigen.
2. B e n u t z u n g s s i c h e r h e i t : Unter Benutzungssicherheit ist zu verstehen, daß es einem Benutzer schwer fällt, vom Modell falsche Aussagen zu erhalten. Das Kriterium der Benutzungssicherheit kann durch eine Modellstruktur erreicht werden, die die Mo-

1 Vgl. zu diesen Kriterien Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

dellerggebnisse auf eine Klasse bedeutsamer Aussagen beschränkt.

3. **P r ü f b a r k e i t :** Der Benutzer eines Modells muß in die Lage versetzt werden, den Modelloutput dahingehend zu überprüfen, daß er Vertrauen in die Ergebnisse des Modells haben kann.
4. **A d a p t i v i t ä t :** Ein Modell ist dann adaptiv, wenn es an neue Inputinformationen angepaßt werden kann. Das Kriterium der Adaptivität gilt primär für die Parameter von Modellen aber auch für deren Struktur.
5. **V o l l s t ä n d i g k e i t :** Das Kriterium der Vollständigkeit besagt, daß Modellstrukturen gefunden werden müssen, die es ermöglichen, viele Phänomene zu berücksichtigen, ohne daß das Modell diese Vielfalt von Phänomenen nicht mehr verarbeiten kann.
6. **K o m m u n i k a t i o n s m ö g l i c h k e i t :** Kommunikationsmöglichkeiten mit dem Modell sind dann gegeben, wenn das Management in der Lage ist, die Modellinputs schnell zu ändern, um die gewünschten Outputs zu erhalten.

Diese und weitere Kriterien gilt es im folgenden unter Beachtung des theoretisch-intersubjektiven Ansatzes dahingehend zu analysieren, ob sie konsistent sind und bei der Entwicklung und Selektion von Marketing-Modellen eine brauchbare Hilfe liefern.

3. Analyse von Kriterien zur Modellselektion und -entwicklung

3.1 Einfachheit und Prüfbarkeit

Das Kriterium der Einfachheit wird von LITTLE¹ an erster Stelle genannt. Es stellt sich die Frage, ob der Begriff der Einfachheit - diese ist dann gegeben, wenn sie das Verständnis des Benutzers erleichtert und wenn bedeutende Phänomene im Modell berücksichtigt und unbedeutende vernachlässigt werden - hinreichend präzise ist, um als erstes Kriterium für die Modellselektion und -entwicklung zu fungieren.

Der Begriff der "Einfachheit" kann unterschiedlich interpretiert werden. Einerseits kann er mit dem Begriff der Kausalität gleichgesetzt werden.² Der Erklärungswert eines Gesetzes ist um so größer, je einfacher dieses Gesetz ist. Kausale Beziehungen sind dann um so strenger, wenn ein Gesetz vorhanden ist und wenn darüber hinaus dieses Gesetz einfach strukturiert ist. Diese Auffassung soll in Anlehnung an POPPER³ als erkenntnis-logischer Einfachheitsbegriff bezeichnet werden. Darüber hinaus unterscheidet POPPER⁴ einen ästhetisch-pragmatischen Einfachheitsbegriff, der dann gegeben ist, wenn Konsens über die Einfachheit bzw. Schwierigkeit bei der Anwendung z.B. eines mathematischen Verfahrens gegeben ist.

Geht man von der Definition des Einfachheitsbegriffs bei LITTLE aus, so läßt sich die hier vertretene Auffassung in die Kategorie des ästhetisch-pragmatischen Einfachheitsbegriffs einordnen. Die Einfachheit eines Modells soll in

1 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

2 Vgl. Schlick, M., Die Kausalität in der gegenwärtigen Physik, in: Naturwissenschaften, 19. Jg., 1931, S. 143ff, hier S. 148.

3 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 98.

4 Vgl. ebenda.

erster Linie das Verständnis seiner Benutzer erleichtern. LITTLE verwendet diesen Begriff ziemlich unkritisch, so als ob es sich von selbst verstünde, was unter Einfachheit zu verstehen ist und warum diesem Kriterium ein bedeutsamer Stellenwert bei der Modellselektion und -entwicklung zukommt. Die grundlegende Kritik an dieser Interpretation des Einfachheitsbegriffs ergibt sich aufgrund der mangelnden intersubjektiven Nachprüfbarkeit, was bei einem Modellbenutzer als einfach zu gelten hat. Für manchen Benutzer ist ein bestimmtes Modell sowohl aufgrund seiner Struktur als auch des bei ihm angewandten mathematischen Verfahrens schwierig, während es für andere Benutzer "trivial" ist. Solange sich - gemessen am Verständnis des Modells - kein einheitlicher Konsens über dessen Einfachheit bzw. Schwierigkeit bei den Marketing-Managern herausbildet, ist dieses Kriterium für die Selektion und Entwicklung von Marketing-Modellen oft fraglich, da es nicht intersubjektiv nachprüfbar ist. Es bietet sich hier nur die Möglichkeit, durch Befragungen über die subjektiv empfundene Einfachheit eine diesbezügliche Einstufung von Marketing-Modellen vorzunehmen. Nur mittels empirischer Untersuchungen kann die erste Komponente des LITTLEschen Einfachheitsbegriffs - die Erleichterung des Verständnisses des Benutzers - "gerettet" werden.

Die zweite Komponente des LITTLEschen Einfachheitsbegriffs knüpft an der Überlegung an, daß Einfachheit dann gegeben ist, wenn bedeutsame Phänomene im Modell berücksichtigt und unbedeutsame vernachlässigt werden.¹ Diese Auffassung kann mit dem Begriff der "einfachsten Beschreibung" gleichgesetzt werden. Ein Modell ist um so einfacher, je mehr es sich auf die wesentlichen Variablen eines Marketing-Phänomens konzentriert, um dem Benutzer Verständnisschwierigkeiten durch unwesentliches Detail zu ersparen.

¹ Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

Bei dieser Begriffsfassung treten wiederum Probleme auf, die darauf zurückzuführen sind, daß nur im Einzelfall bestimmt werden kann, was "wesentliche Variablen eines Marketing-Phänomens" und was "vernachlässigungswerte Phänomene" sind. Um das Kriterium der "einfachsten Beschreibung" zu intersubjektivieren, bietet sich wiederum die subjektive Einstufung von Modellen durch Relevanzziffern an. Eine diesbezügliche empirische Untersuchung müßte darüber hinaus dahingehend erweitert werden, was bei Marketing-Modellen als wesentlicher und was als unwesentlicher Modellbestandteil anzusehen ist. Erst wenn mit dieser Untersuchung ein allgemeiner Konsens über berücksichtigungswürdige und vernachlässigungswerte Modellbestandteile sich herausbildet, kann der zweiten Komponente des LITTLEschen Einfachheitsbegriffs Aussagefähigkeit beigemessen werden.

In diesem Zusammenhang ist POPPER¹ zuzustimmen, der den ästhetisch-pragmatischen Einfachheitsbegriff aus logischen Gründen ablehnt und dazu feststellt: "Ähnlich steht es, wenn man sagt, eine Aufgabe sei mit "einfacheren Mitteln" lösbar als eine andere, und damit etwa meint, daß ihre Lösung weniger Vorkenntnisse voraussetzt. In allen solchen Fällen kann man das Wort "einfach" ... leicht eliminieren, da seine Verwendung eine außerlogische ist."²

Mit diesem Zwischenergebnis soll keineswegs in Abrede gestellt werden, daß es für den Benutzer "einfache" und "schwierige" Modelle gibt. Vielmehr sollte bewiesen werden, daß ohne empirische Untersuchungen der ästhetisch-pragmatische Einfachheitsbegriff derart mit Unschärfe behaftet ist, daß sich auf ihn in nicht-empirischen Situationen keine Aussagen gründen lassen, die Aufschluß über die Einfachheit bzw. Schwierigkeit eines Modells geben. Aus diesem Grund scheint es angebracht zu sein, den erkenntnislogi-

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 98.

2 Ebenda.

schen Einfachheitsbegriff näher zu analysieren, um davon ausgehend, das LITTLEsche Postulat der Einfachheit näher präzisieren zu können.

Der erkenntnislogische Einfachheitsbegriff knüpft an der Überlegung an, daß jeder Forscher, dem es gelungen ist, eine Beobachtungsreihe durch eine sehr einfache Formel oder Aussage zu beschreiben, davon überzeugt ist, ein Gesetz gefunden zu haben.¹ Je strenger die Kausalität im Sinne einer Ursache-Wirkungs-Beziehung ist, um so einfacher wird aller Wahrscheinlichkeit nach dieses Gesetz beschaffen sein. Dieser erkenntnislogischen Begriffsfassung der Einfachheit stehen aber eine Reihe von Schwierigkeiten gegenüber. So stellt SCHLICK² fest, "daß Einfachheit offenbar ein ganz relativer und unscharfer Begriff ist, so daß eine strenge Definition der Kausalität nicht erreicht wird und Gesetz und Zufall sich nicht voneinander unterscheiden lassen." Insbesondere hat die "Einfachheit" bei mathematischen Marketing-Modellen als Selektionskriterium eine Bedeutung, wenn solche Modelle bevorzugt zur Anwendung gelangen sollen, die es gestatten, auf möglichst einfache Weise zu einer Lösung zu gelangen. Frage ist, ob das Postulat der Einfachheit Charakteristikum des Modells, der in das Modell eingehenden Variablen und Parameter oder beides ist.

Mehrere Ansätze versuchen den Einfachheitsbegriff mittels Wahrscheinlichkeiten zu definieren.³ Werden von einigen ursprünglich vorhandenen Hypothesen diejenigen ausgeschlossen, die mit dem im Modell zu berücksichtigenden Daten unverträglich sind, so bleiben einige gleichstarke Hypothesen übrig, die mit dem Datenmaterial kompatibel

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 98.

2 Schlick, M., Die Kausalität in der gegenwärtigen Physik, a.a.O., S. 148.

3 Vgl. z.B. Carnap, R. / Stegmüller, W., Induktive Logik und Wahrscheinlichkeit, Wien 1959, S. 23ff und Behnke, H. et al., Mathematik, Bd. 2, Frankfurt 1966, S. 347.

sind. Der Marketing-Manager wird nun mit der Hypothese arbeiten, die unter den gemachten Voraussetzungen am wahrscheinlichsten ist. Beispielsweise führt dieses bei der Entwicklung von Marketing-Modellen dazu, bestimmte Größen zu vernachlässigen, weil deren Eintritt unwahrscheinlich ist. Es wird somit ein Maß für die Wahrscheinlichkeit der unter den getroffenen Voraussetzungen aufgestellten Hypothese benötigt. Derartige Wahrscheinlichkeitsmaße lassen sich nicht zahlenmäßig-generalisierend festlegen, jedoch wird man in konkreten Situationen gewisse Prinzipien angeben können, denen derartige Maßfunktionen im allgemeinen genügen werden.¹

An dieser Stelle sei auf das von HERMES² vertretene Einfachheitsmaß verwiesen, das bei gleichen gewonnenen Erfahrungen die sprachlich kürzere Hypothese der sprachlich längeren vorzieht, da diese einfacher ist. Da das Maß der sprachlichen Kürze der Hypothese stark von der Art der verwendeten Sprache abhängt, ist bei mathematisch formulierten Hypothesen, um die es hier ausschließlich geht, die sprachliche Kürze ein unscharfes Unterscheidungskriterium.

Das Einfachheitsproblem stellt sich in besonders prägnanter Weise immer dann, wenn es darum geht, eine Reihe von Beobachtungspunkten durch eine möglichst einfache Kurve miteinander zu verbinden.³ So läßt sich z.B. durch eine endliche Zahl von beobachteten Absatzwerten eine unendliche Kurvenschar legen, wobei sich die einzelnen Kurven durch ihre Funktionstypen und Wertigkeiten unterscheiden. Letztlich sind alle Versuche, den Einfachheitsbegriff auf den Funktionstyp und die Wertigkeit zurückzuführen, zum Scheitern verurteilt, da nicht einzusehen ist, warum

1 Vgl. dazu Kneale, W., Probability and Induction, Oxford 1949, S. 229f und Carnap, R. / Stegmüller, W., Induktive Logik und Wahrscheinlichkeit, a.a.O., S. 155ff.

2 Vgl. Hermes, H., Aufzählbarkeit - Entscheidbarkeit - Berechenbarkeit, Berlin - Göttingen - Heidelberg 1961, S. 25ff und S. 46ff.

3 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 100.

lineare Funktionen einfacher als z.B. e-Funktionen sein sollen. Ein derartiger Einfachheitsbegriff hat überwiegend pragmatischen Charakter.

Interessanter ist der Versuch, die Einfachheit einer Funktion durch ihren Freiheitsgrad zu definieren.¹ Dieser Gedankengang läßt sich näher präzisieren, wenn man die Anzahl der Freiheitsgrade näher untersucht. Bei einem Gleichungssystem mit z.B. sechs Freiheitsgraden benötigt man die doppelte Anzahl an Variablen, um das System lösen zu können. In diesem Sinn ist die Einfachheit einer Funktion dann "groß", wenn die Anzahl ihrer frei anpaßbaren Parameter ebenfalls "groß" ist. Das Ergebnis kann hier nur lauten, daß der Einfachheitsbegriff nicht über Modellrelationen ableitbar ist.

Unter Berücksichtigung des Falsifizierungsgrades und seiner Identifizierung mit dem Einfachheitsbegriff läßt sich das bislang ungelöste Problem der Einfachheit einer Lösung zugänglich machen.² Einfach strukturierte Theorien lassen sich leichter falsifizieren als komplexere Theorien.³ "Einfache Sätze sind ... deshalb höher zu bewerten als weniger einfachere, weil sie mehr sagen, weil ihr empirischer Gehalt größer ist, weil sie besser überprüfbar sind."⁴ Das Selektionskriterium der Einfachheit läßt sich somit auf das der Überprüfbarkeit zurückführen. Letzteres entscheidet somit darüber, ob ein Modell - verstanden als Menge nomologischer Hypothesen - einfach ist oder nicht. Die hier genannten Gesichtspunkte zum Begriff der Einfachheit gestatten es, folgende Thesen über die von LITTLE postulierte Einfachheit von Marketing-Modellen aufzustellen:

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 100.

2 Vgl. ebenda, S. 100f.

3 Vgl. ebenda, S. 101.

4 Ebenda, S. 103 (Im Original teilweise kursiv).

1. Das Problem der Einfachheit (Überprüfbarkeit) ist nicht ein Problem der in einem Modell angewandten Funktionstypen, sondern ein Problem der Datenbeschaffung, um diese Funktionen testen zu können.
2. Da ein Modell aus mehreren miteinander verbundenen Funktionen besteht, ist dessen Einfachheit um so eher dann gegeben, wenn die Überprüfbarkeit von Modellzwischenenergebnissen gegeben ist.

Die Diskussion des Einfachheitsbegriffs führt zu der generellen Schlußfolgerung, daß sowohl die ästhetisch-pragmatische als auch die erkenntnislogische Interpretation des Einfachheitsbegriffs nur dann sinnvolle Kriterien für die Selektion und Entwicklung von Marketing-Modellen sind, wenn sie mit dem Kriterium der Prüfbarkeit gleichgesetzt werden. Erst durch empirische Untersuchungen von Modellen, bei denen ein allgemeiner Konsens über die Einfachheit und Prüfbarkeit der Modelle erzielt wird, kann "Einfachheit" zu einem sinnvollen - d.h. intersubjektiv nachprüfbaren - Selektions- und Entwicklungskriterium für Marketing-Modelle werden. Darüber hinaus überschneiden sich bei der hier vorgenommenen Begriffsexplikation des Einfachheitsbegriffs die von LITTLE¹ angeführten Kriterien der Einfachheit und Prüfbarkeit. Aus Gründen der möglichst operationalen Begriffsfassung der Einfachheit scheint es nicht angebracht zu sein, eine Trennung zwischen Einfachheit und Prüfbarkeit vorzunehmen, da die Prüfbarkeit nach der hier vertretenen Auffassung nur ein operationales Begriffsexplikat der Einfachheit ist.

¹ Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

3.2 Prognostische Relevanz

Ein Modell ist um so eher für die Aufnahme in die Modellbank eines Marketing-Informationssystems geeignet, je mehr es mittels dieses Modells gelingt, künftige Werte aus vergangenen Werten zu erklären. Die Überlegung lautet, daß sich die dazu erforderlichen Informationen beschaffen lassen, dann Funktionen spezifiziert werden können, diese Funktionen prognostisch relevant sind, so daß sich die Zukunft mathematisch fassen läßt.

Die prognostische Relevanz ist aller Wahrscheinlichkeit nach entscheidend für die Gültigkeit eines Marketing-Modells. Wenn die Marketing-Manager mittels eines Prognosemodells in der Lage sind, das Verhalten von bedeutsamen Modellvariablen vorherzusagen, wird ihr Vertrauen in die prognostischen Fähigkeiten dieses Modells anwachsen, so daß sie es vielleicht anwenden werden. In diesem Fall wäre der Einwand von LITTLE¹ ausgeräumt, daß das grundsätzliche Problem bei Marketing-Modellen darin liegt, daß sie vom Management nicht angewendet werden.

Das Kriterium der prognostischen Relevanz wird von LITTLE implizit in seinen Kriterien der Benutzungssicherheit und Prüfbarkeit berücksichtigt.² Die Benutzungssicherheit - d.h. die Schwierigkeit der Generierung falscher Modellaussagen durch den Benutzer - wird durch Modellstrukturen erzielt, die die Modellergebnisse auf eine Klasse von bedeutsamen Modellaussagen beschränken.³ Darüber hinaus ist im Zusammenhang mit der Prüfbarkeit von Modellaussagen der Gesichtspunkt der Genauigkeit dieser Aussagen von entscheidender Bedeutung. Problem ist, ob und inwieweit sich diese Kriterien dergestalt operationalisieren lassen, daß sowohl dem Modellkonstrukteur als auch dem Marketing-Management Entscheidungshilfen bei der Selektion und Entwick-

1 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-466.

2 Vgl. ebenda, S. B-470.

3 Vgl. ebenda.

lung von Marketing-Modellen an die Hand gegeben werden.

Sowohl Modellkonstrukteure als auch Entscheidungsträger stehen bei der Aufgabe der Selektion und Entwicklung zweckmäßiger Marketing-Modelle vor einer äußerst komplexen Aufgabe. Die Bewältigung dieser Aufgabe übersteigt die Informationsverarbeitungsfähigkeiten einzelner Modellkonstrukteure und Entscheidungsträger bei weitem. Aus diesem Grund wird oft "Zuflucht" zu den Wahrscheinlichkeitsgesetzen genommen, die angesichts des unvollkommenen Informationsstandes des Entscheidungsträgers gute Dienste leisten.¹ Die Anwendung des LAPLACESchen Wahrscheinlichkeitskriteriums² erfordert aber vollständige und genaue Informationen über den Zustand in einem vorgegebenen Zeitpunkt. Es erhebt sich die Frage, ob ein derartiger Zustand überhaupt möglich ist. Es ist einsichtig, daß absolute Genauigkeit nicht zu erreichen ist.³ Der Genauigkeitsgrad einer Beobachtung läßt sich durch die Verbesserung der Beobachtungstechnik⁴ stark vergrößern. Daraus ergibt sich die Konsequenz, daß bei Verfeinerungen der Beobachtungstechnik die Prognoseinformationen immer zuverlässiger werden. Es ist letztlich die stillschweigende Voraussetzung

-
- 1 Vgl. zu den Wahrscheinlichkeitsüberlegungen in Zusammenhang mit unvollkommenem Informationsstand Wittmann, W., Unternehmung und unvollkommene Information - Unternehmerische Voraussicht - Ungewißheit und Planung, Köln und Opladen 1959, S. 79ff.
 - 2 Nach der Laplaceschen Wahrscheinlichkeitsinterpretation sind sämtliche möglichen Ereignisse gleichwahrscheinlich, so daß bei unvollständigen und ungenauen Informationen über den Zustand eines Ereignisses zu einem bestimmten Zeitpunkt kein Kriterium für die Vorziehenswürdigkeit eines Ereignisses besteht. "Aus der Unkenntnis der Wahrscheinlichkeiten ist nicht zu schließen, daß man sich so verhalten müßte, als ob alle möglichen Ereignisse gleichwahrscheinlich wären und schon gar nicht, daß alle diese Ereignisse gleichwahrscheinlich sind." Krelle, W., Preistheorie, Tübingen - Zürich 1968, S. 99.
 - 3 Vgl. dazu die Ausführungen über Scheingenauigkeit bei Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 61-68.
 - 4 Vgl. Ackoff, R.L., Scientific Method, a.a.O., S. 209.

dieser Vorgehensweise, daß man nicht allzu schnell auf Grenzen stößt, die die Genauigkeit einengen.

Bei Beobachtungen im Marketing hängt aber der zu realisierende Beobachtungsprozeß von den Ausgangsbedingungen ab. Dies ist auf die deduktive Erklärungsaufgabe einer Menge von Marketing-Phänomenen zurückzuführen, wobei es sich hier um die kausale Erklärung handelt. Einen Marketing-Prozeß oder ein Marketing-Phänomen kausal zu erklären, bedeutet einen beschreibenden Satz aus Quasi-Gesetzen und Randbedingungen deduktiv abzuleiten.¹ Es stellt sich somit die Aufgabe, ein Modell zu finden, das bei empirischer Überprüfung die gleichen Daten liefert wie dieser theoretische Satz. Wenn empirische Daten aber mit theoretischen Modellaussagen verglichen werden, erfolgt die implizite Voraussetzung, daß Marketing-Phänomene reproduzierbar sind. Bei Wiederholung derselben empirischen Untersuchung unter gleichen Bedingungen müssen sich identische oder annähernd identische Beobachtungsdaten einstellen. Dieses ist die Voraussetzung, derzufolge eine "vernünftige" Struktur des zu analysierenden Phänomens existiert. In diesem Zusammenhang stellt LITTLE² fest, daß die Benutzungssicherheit eines Modells durch eine Struktur erreicht wird, die die Modellaussagen auf eine Klasse bedeutsamer Aussagen beschränkt. Daraus ist zu folgern, daß dieses LITTLEsche Postulat nur dann zu erreichen ist, wenn sich das Modell für die betreffende Fragestellung eignet, wenn die Modellprämissen realistisch sind und die dem Modell zugrundeliegenden nomologischen Hypothesen prognostisch relevant sind. Dies führt zu dem Resultat, daß aus Einzelfällen abgeleitete Schlußfolgerungen und die sich aus diesen Schlußfolgerungen zu entwickelnden Erfahrungsregeln für Einzelfälle zwangsläufig zu Modellen führen müssen, deren prognostische

1 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 31.

2 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

Relevanz gering ist.¹ Ein typisches Beispiel dafür sind Modelle, die mit Wahrscheinlichkeiten arbeiten, obwohl die Voraussetzungen des zentralen Grenzwertsatzes der Wahrscheinlichkeit nicht gegeben sind, so daß die prognostische Relevanz derartiger Modelle - abgesehen vom Einzelfall - gering ist.

Ein weiterer Grund für nur geringe prognostische Relevanz von Modellen ist auf die Tatsache zurückzuführen, daß die Modelle oft nicht so beschaffen sind, daß der Modellausgangszustand sämtliche möglichen künftigen Zustände determiniert. Es ist zu beachten, daß das dieser Aussage zugrundeliegende Kausalitätsprinzip nicht so verstanden werden darf, daß eine Ursache die Wirkung hervorbringt. Vielmehr besteht nur dann Kausalität, wenn zwei Ereignisse U und W in der Beziehung von Ursache und Wirkung stehen, sofern sie drei Bedingungen erfüllen, nämlich

- wenn sie in Zeit und Raum aufeinander folgen,
- wenn das Ereignis U dem Ereignis W zeitlich vorausgeht
- und
- wenn das Ereignis W auf das Ereignis U mit großer Wahrscheinlichkeit folgt.

Im Zusammenhang mit diesen Bedingungen² muß ein viertes Element, welches der Kausalität anzuhaften erscheint, die Unvermeidlichkeit oder der Zwang, als nicht charakteristisch angesehen und aus der Betrachtung ausgeschlossen werden. Es zeigt sich, daß die hier gemeinte Kausalität - gleichgültig

1 Bei derartigen Modellen ist die Bedingung i.d.R. nicht erfüllt, daß sie mindestens eine Aussage enthalten müssen, die einen theoretischen Begriff enthält, mit dessen Hilfe Prognosen über zukünftige Ereignisse abgeleitet werden können, die ohne diese Aussage nicht möglich sind. Vgl. Stegmüller, W., Hauptströmungen der Gegenwartsphilosophie, a.a.O., S. 456.

2 Vgl. zu einer vereinfachten Systematisierung der Kausalzusammenhänge Förster, E. / Egermayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 14.

ob in populärer Version oder im Licht kritischer Überprüfung verstanden - auf Marketing-Modelle nur beschränkt anwendbar ist. Es wäre oft völlig willkürlich, irgendeinen "Modellbestandteil" herauszugreifen und ihn als Ursache zu bezeichnen und ebenso eine andere Konfiguration als Wirkung zu definieren. In der Symbolsprache mathematischer Modelle kommen derartige Kennzeichnungen auch nicht vor. An ihre Stelle tritt der Begriff der funktionalen Abhängigkeit zwischen bestimmten Variablen und Parametern.¹ Wichtig ist, daß eine solche zwischen zwei oder mehreren Variablen bestehende funktionale Abhängigkeit keineswegs die Existenz einer kausalen Beziehung zwischen den Größen bedingt, sondern lediglich eine Zuordnung zwischen ihren Werten herstellt.²

Betrachtet man in einer Marktreaktionsfunktion die funktionale Abhängigkeit zwischen erklärter Variable und erklärenden Variablen, so läßt sich diese Funktion nach jeder der erklärenden Variablen auflösen und liefert die betreffende Variable als Funktion der anderen Variablen, ohne daß angenommen zu werden braucht, daß etwa eine Änderung der erklärten Variable die "Ursache" für die Änderung einer oder sämtlicher erklärenden Variablen ist und umgekehrt. Nur bei Geltung der oben angeführten Bedingungen kann für den Einzelfall Kausalität angenommen werden. Wenn z.B. der Preis erhöht wird, so geht die nachgefragte Menge zurück. Die Preiserhöhung geht der Verringerung der Nachfrage zeitlich voraus, wobei diese unmittelbar auf die Preiserhöhung folgt. Darüber hinaus muß sich aufgrund der Preiserhöhung mit großer Wahrscheinlichkeit eine Verringerung der nachgefragten Menge ergeben. Ob aber tatsächlich die Preiserhöhung Ursache für die Nachfrageverringering

1 "Sobald es gelingt, die Elemente und Ereignisse durch meßbare Größen zu charakterisieren, ... läßt sich die Abhängigkeit der Elemente voneinander durch den Funktionsbegriff viel vollständiger und präziser darstellen als durch so wenig bestimmte Begriffe wie Ursache und Wirkung." Mach, E., Erkenntnis und Irrtum, 4. Aufl., Leipzig 1920, S. 278.

2 Vgl. ebenda, S. 278ff.

ist, bleibt einzelfallabhängig.

Die Postulierung von funktionalen Zusammenhängen eröffnet die Möglichkeit, insbesondere bei dynamischen Beziehungen, die Zukunft aus der Gegenwart abzuleiten und Prognosen zu erstellen.¹ Es ist Kennzeichen sämtlicher im Marketing angewandter Prognoseverfahren, daß es durch die Annahme funktionaler Zusammenhänge gelingt, aus gegenwärtigem Wissen zukünftiges Wissen zu gewinnen. Bei dynamischen Beziehungen läßt sich durch Auflösung nach einem bestimmten Zeitindex vom gegenwärtigen Zustand auf den Zustand des Systems in der Vergangenheit - und wenn Stetigkeit der Entwicklung unterstellt wird - auch auf zukünftige Systemzustände schließen.² Die Stetigkeit der Entwicklung ist durch die drei angeführten Kausalitätsbedingungen, insbesondere durch die dritte Bedingung, gegeben. Dies hat seinen Grund darin, daß dynamische Gleichungen bei der Umkehr der Zeitrichtung ihre Form beibehalten, d.h. gegenüber der Ersetzung von t durch $-t$ invariant sind.

In mathematischen Gleichungen von Marketing-Modellen wird zwar bestimmten Phänomenen der Charakter von Ursachen und den damit zusammenhängenden Erscheinungen zugeschrieben, doch dann ist das jeweilige Phänomen kein Ereignis, da nur Ereignissen eine Kausalbeziehung zugeschrieben werden kann. Die Bezeichnung von "Ursache" und "Wirkung" ist in Marketing-Modellen aufgrund der angestrebten Erklärung notwendig. Im mathematischen Sinn ist in Modellgleichungen aber nichts vorhanden, was mit der Auffassung von "Ursache" und "Wirkung" in Zusammenhang zu bringen ist. Die Gleichungen beschreiben zwar, was sich unter bestimmten Bedingungen ereignet, aber aus ihnen läßt sich keine Ursache entnehmen,

1 Zum Zusammenhang zwischen Kausalitätsprinzip und Prognosemodellen vgl. Stachowiak, H., Über kausale, konditionale und strukturelle Erklärungsmodelle, in: *Philosophia Naturalis*, Bd. IV, H. 4, 1957, S. 403ff und Popper, K.R., *Logik der Forschung*, a.a.O., S. 31ff.

2 Dieses geschieht durch Bildung der inversen Funktionen.

w a r u m sich die Phänomene gerade so und nicht anders ereignen.

Mit dem Aufkommen der quantitativen Modellanalyse im Marketing gewann auch dieses Problem der Kausalität einen neuen Aspekt: Die Frage nach dem "warum" wurde durch die Frage nach dem "wie" verdrängt.¹ Durch diese Vorgehensweise soll die Existenz einer Kausalbeziehung nicht in Abrede gestellt werden. Es existiert ein Kausalzusammenhang, nur nicht in der Version als Beziehung zwischen Ereignissen.

Die Gesetze im Marketing sind so beschaffen, daß sich aus der Kenntnis des Zustands eines geschlossenen Systems zu einem bestimmten Zeitpunkt der Zustand des Systems zu einem anderen Zeitpunkt nicht bestimmt werden kann, so daß der gesamte Ablauf bzw. die sich vollziehenden Prozesse aufgrund der Stochastizität von Marketing-Gesetzen nicht im voraus angegeben werden können. Angesichts dieser Tatsache muß das Kausalitätsprinzip anders gefaßt werden. Die Prozesse im Marketing werden vor allem durch Kontinguität bestimmt. Was an einem bestimmten Zeitpunkt stattfindet, hängt primär davon ab, was sich "kurz vorher" vollzogen hat. Dieses Kontinguitätsprinzip ergibt sich aufgrund der Tatsache, daß die Marketing-Manager zumindest mittelfristig mit tendenziell stetigen Entwicklungen rechnen und somit die Möglichkeit zur Prognose besitzen.

Eine derart kontinguitive Kennzeichnung des Kausalitätsprinzips muß notwendigerweise umrißhaft sein. Mathematisch argumentiert heißt dies, daß an die Stelle gewöhnlicher partielle Differentialgleichungen, an die Stelle

¹ Zum Zusammenhang zwischen betriebswirtschaftlichen Experimenten und Kausalfaktoren vgl. Cochran, W.C., The Design of Experiments, in: Flagle, C.D. / Huggins, W.H. / Roy, R.H., Hrsg., Operations Research and Systems Engineering, Baltimore 1960, S. 508ff, ferner Stachowiak, H., Denken und Erkennen im kybernetischen Modell, a.a.O., S. 143.

von Anfangsbedingungen Randbedingungen treten.¹ Die im Marketing feststellbaren Gesetze sind aber nur Quasi-Gesetze, die es nicht gestatten, in präziser Weise zu prognostizieren. Die prognostische Relevanz der in der Marketing-Theorie feststellbaren ontologischen und nomologischen Hypothesen ist zu gering, um von gesetzmäßigen Entwicklungen zu sprechen. Somit bleibt nur die Möglichkeit, das Kausalitätsprinzip als eine Voraussetzung für die Möglichkeit von Erfahrung zu deuten, so daß es von der Erfahrung selbst gar nicht mehr in Zweifel gezogen werden kann.

Dies führt zum Schluß, die Marketing-Theorie verstärkt auf "erfahrungswissenschaftlichen Theorien" basieren zu lassen, die sich als bewährt erwiesen haben und von einer Vielzahl von Erfahrungen nicht oder nur bedingt falsifiziert werden können. Läßt man erfahrungswissenschaftliche Gesetze von beliebiger Form im Marketing zu, so ist es immer möglich, eine gegebene Ereigniskette, und mag sie noch so bizarr und unregelmäßig sein, durch eine mathematische Funktion darzustellen. Mathematisch ausgedrückt heißt dies: Jede unregelmäßige Kurve kann durch eine Funktion abgebildet werden, vorausgesetzt, die Kurve ist innerhalb eines gewissen Intervalls stetig und besitzt dort höchstens endlich viele Unstetigkeitsstellen (Fourierreihe).²

Ein Beispiel soll diesen Zusammenhang näher verdeutlichen. Wird aufgrund von Absatzmengen-Beobachtungen der Vergangenheit eine endliche Anzahl von Beobachtungswerten gewonnen, so gibt es zumindest eine ganze rationale Funktion vom Grade $k-1$, nämlich $f(x) = a_1 + a_2x + \dots + a_mx^{k-1}$, die die

1 "Wir finden also zwei verschiedene Arten von Sätzen, die erst gemeinsam die vollständige 'kausale Erklärung' liefern: (1) allgemeine Sätze... und (2) besondere Sätze, die nur für den betreffenden Fall gelten - die 'Randbedingungen'." Popper, K.R., Logik der Forschung, a.a.O., S. 32.

2 Vgl. dazu v. Mangoldt, H. / Knopp, K., Einführung in die höhere Mathematik, Bd. 3, 9. Aufl., Stuttgart 1948, S. 502ff.

beobachteten Absatzmengen völlig beschreibt. Die darüber hinaus noch bestehende Anzahl Polynome höherer Ordnung (Fourierreihe) weist gegenüber der ganzen rationalen Funktion keine "Besonderheit" auf, daß sie dieser rationalen Funktion vorgezogen werden könnte. Das Kausalitätsprinzip wäre in diesem Fall ziemlich inhaltsleer.

Eine Möglichkeit, sich diesem Dilemma zu entziehen, besteht darin, folgendes von einem Quasi-Gesetz im Marketing zu verlangen: Es muß so beschaffen sein, daß auf seiner Grundlage Vorhersagen möglich sind.¹ Darüber hinaus soll das Gesetz so beschaffen sein, daß es gelingt, Daten in Beziehung zueinander zu setzen und zu erklären. Ein erfahrungswissenschaftliches Gesetz muß mit allen Daten verträglich sein, und falls sich mehrere Möglichkeiten bieten, gibt im allgemeinen die Einfachheit den Ausschlag. Das Seltsame und recht Überraschende daran ist, daß Einfachheit und Prognostizierbarkeit oft zusammen auftreten, eine Tatsache, die als Rationalität ausgelegt werden kann.³ Diese Rationalität und der Umstand, daß "Ursache" und "Wirkung" oft mit Regelmäßigkeit aufeinander folgen, ist für das Marketing von erheblicher Bedeutung. Das Kausalgesetz ist hier gleichbedeutend mit der geläufigen Erkenntnis, daß man zwischen zwei Ereignissen, einem beliebig herausgegriffenen und einem anderen, das ihm vorausgeht, immer wieder gleiche oder annähernd gleiche Prozesse fest-

1 Vgl. Chmielewicz, K., Forschungskonzeptionen der Wirtschaftswissenschaft - Zur Problematik einer entscheidungstheoretischen und normativen Wirtschaftslehre, Stuttgart 1970, S. 16.

2 Vgl. dazu die Ausführungen über den ästhetisch-pragmatischen und den erkenntnislogischen Einfachheitsbegriff im Punkt 3.1. dieses Abschnitts.

3 Vgl. Popper, K.R., Logik der Forschung, a.a.O., S. 100ff.

stellen kann.¹ Diese können dann auf ihre empirische Gesetzmäßigkeit hin überprüft werden.²

Die dabei vorzunehmende Analyse besteht darin, den Ablauf der wirklich stattgefundenen Ereignisse (gedanklich) in einzelne Elementarprozesse zu zerlegen, die häufig wiederkehren und gemessen werden können. Insbesondere ist die system- und entscheidungsorientierte Marketing-Theorie darauf ausgerichtet, die Realität zu begreifen, indem sie diese in ihre Bestandteile zerlegt, dann nach Gesetzen sucht, die zwischen den Komponenten bestehen und schließlich diese Realität aus diesen Elementen wieder zusammensetzt. Durch diesen Analyse- und Syntheseprozess gelingt es aber der Marketing-Theorie fast nur ausschließlich, Quasi-Gesetze abzuleiten, die es nicht gestatten, beliebig herausgegriffene Phänomene miteinander in Beziehung zu setzen.

Die Marketing-Theorie befaßt sich mit vereinfachten Beziehungen zwischen Systemelementen. Sobald man diesen Elementen mathematische Größen zuordnet, kann es vorkommen, daß man auf funktionale Abhängigkeiten zwischen diesen stößt, die dann zur Entdeckung von Quasi-Gesetzen führen. Das Quasi-Gesetz tritt dann an die Stelle von Ursache-Wirkungs-Beziehungen und es erübrigt sich, derartige Beziehungen genauer zu untersuchen. Ob Kausalität besteht oder nicht, ist dann an die Bedingung geknüpft, daß Beobachtungen und Messungen mit sehr hoher Genauigkeit vorgenommen werden können. Sollte diese Bedingung aus dem einen oder anderen Grund unerfüllbar sein, entfällt die Möglichkeit zu prognostizieren. Das bedeutet aber, daß das Kausali-

1 Dieses gilt nur unter der Bedingung, daß bei "sicheren" Gesetzen die Randbedingungen nicht verschoben werden, so daß keine Aussagen mehr möglich sind. Vgl. Seiffert, H., Einführung in die Wissenschaftstheorie, a.a.O., S. 177.

2 Vgl. Kamlah, W. / Lorenzen, P., Logische Propädeutik oder Vorschule des vernünftigen Redens, a.a.O., S. 139.

tätsprinzip am Kriterium der Vorhersagbarkeit gemessen, seine Gültigkeit verliert. Zwar wird es nicht im üblichen Sinne falsch, aber es kann strenggenommen nicht mehr als gültig angesehen werden, weil die in ihm steckenden Voraussetzungen nicht zu realisieren sind.

Damit ist das LITTLEsche Postulat der Vermeidung von falschen Modellaussagen¹ durch den Benutzer aufgrund der Beschränkung dieser Aussagen auf diejenigen, die prognostisch relevant sind, in Frage gestellt. Wenn dieses Postulat zutreffen soll, müssen die Modellrelationen quasisgesetzmäßig sein und diese Quasi-Gesetzmäßigkeiten von der Marketingforschung mit sehr hoher Genauigkeit festgestellt werden können. Diese Bedingungen sind aber im Marketing so gut wie niemals gegeben. So kann sich bereits aufgrund einer geringfügigen Änderung eines Ausgangszustandes ein abweichendes Modellergebnis einstellen. Darüber hinaus werden Prognosen² noch dadurch erschwert, daß der Ausgangszustand der Untersuchung nicht mit großer Genauigkeit bekannt ist. Der geringste Fehler bei dem in einem Prognosemodell verwendeten Daten würde zu einem fehlerhaften Output führen, so daß es im Gegensatz zum LITTLEschen Postulat es einem Benutzer nicht schwerfällt, vom Modell falsche Aussagen zu erhalten.

Es soll hier aufgrund der Forderung nach hoher Genauigkeit nicht zu dem Schluß gekommen werden, daß Marketing-

1 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

2 Der wissenschaftliche Wert von Prognosen wird darüber hinaus angezweifelt: "Wenn Popper also sagt: 'Vom Standpunkt der praktischen Nützlichkeit der Wissenschaft ist die Bedeutung wissenschaftlicher Prognosen ziemlich klar' [Popper, K.R., Das Elend des Historizismus, Tübingen 1965, S. 35], so kann das nur so viel heißen: Zwar brauchen wir die wissenschaftlichen Methoden..., um solche Prognosen machen zu können; aber die Prognosen selbst sind keine wissenschaftlichen Aussagen und haben daher für die Wissenschaft selbst keinerlei Bedeutung." Seiffert, H., Einführung in die Wissenschaftstheorie, a.a.O., S. 184.

Modelle keine sicheren und damit prognostisch-relevanten Aussagen liefern. Es soll hier nur festgestellt werden, daß gewisse Beobachtungsmängel vorliegen, die bei einer oberflächlichen Analyse nicht wahrgenommen werden, die aber das Ergebnis nichtsdestoweniger im gewohnten Sinn kausal bestimmen. Es ist problematisch, Prozesse im Marketing so zu ordnen, daß sich Kausalketten bilden lassen, ja es ist oft unmöglich, sie in einem einheitlichen Bezugsrahmen zusammenfassend darzustellen. Die übliche Forderung nach Anschaulichkeit der Kausalordnung, die LITTLE anscheinend vorschwebt¹, ist selten zu erfüllen. Die Forderung nach Anschaulichkeit der Kausalordnung von gewissen Ereignissen im Marketing führt zu bedeutsamen Konsequenzen bezüglich des Kausalitätsprinzips. Wenn der Gedanke der lückenlosen Beschreibung aufgegeben werden muß, dann läßt sich das Kausalitätsprinzip nicht mehr halten, denn es steht und fällt in seiner wissenschaftlichen Form mit einer solchen Beschreibungsmöglichkeit. In dem Augenblick, wo sich kontinguitive Beschreibung als unmöglich erweist, wird dem Kausalitätsprinzip jegliche Grundlage entzogen. Kontinguitive Beschreibung ist gerade die notwendige Bedingung, welche Kausalität erst ermöglicht. Die Verletzung dieser Bedingung hat eine Verletzung der Kausalität zur Folge.

Im folgenden soll näher untersucht werden, warum eine solche Beschreibung auf Schwierigkeiten stößt. Werden von der Marketing-Forschung z.B. zwei Beobachtungen vorgenommen, die in sehr kurzer Zeit aufeinander folgen, so müssen sie als zwei voneinander getrennte Ereignisse angesehen werden, die nicht zusammenhängen und folglich voneinander unabhängig sind. Es ist unmöglich, sie widerspruchsfrei zusammenzufassen. Darüber hinaus ist es problematisch, die Ergebnisse dieser Beobachtungen mit Sicherheit vorauszusagen. Was man voraussagen kann, ist lediglich die Wahrscheinlich-

¹ Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-483.

keit des Eintritts bestimmter Ereignisse.¹ Ebenso ist es aufgrund der beiden Beobachtungen unmöglich festzustellen, was sich zwischen den beiden Beobachtungszeitpunkten "abgespielt" hat. Mit anderen Worten enthält jede Untersuchung von Ereignissen "Lücken", die sich durch die vorgenommenen Beobachtungen nicht schließen lassen. Mathematisch kann diese "Lücke" nur durch partielle Differentialgleichungen geschlossen werden.² Jede Vorstellung, die von einem benutzungssicheren Modell im LITTLEschen Sinn entworfen wird, muß gemessen an den hier vorgetragenen Gesichtspunkten eo ipso fehlerhaft sein.³

Um die Aussage von LITTLE zu "retten", daß prognostische Relevanz nur über eine spezifische Modellstruktur zu erreichen ist, die die Modellaussagen auf Klassen von prognostisch-relevanten Aussagen beschränkt, bedarf es einer statistischen Marketing-Theorie. Diese Theorie hätte sich nicht mehr mit Einzelereignissen im Marketing, sondern mit dem Verhalten großer Kollektive zu befassen. Sie besitzt folglich eine andere Struktur als die bislang betriebene "exakte" Theorie. Die mittels ihr abgeleiteten Aussagen und entwickelten Modelle sind wahrscheinlichkeitstheoretischer Natur, die das durchschnittliche Verhalten dieser Kollektive betreffen. Dabei ist es völlig willkürlich, ob die "wirklichen Gesetze", also jene, die das Verhalten dieser Kollektive bestimmen, kausalen Charakter haben oder nicht. Jeder Elementarprozeß kann sich entwickeln wie der Zufall es will, und doch werden sich infolge der großen Anzahl von Prozessen im Marketing alle derartigen Schwankungen schließlich herausmitteln und zu dem dem Be-

1 Vgl. Popper, K.R., Das Elend des Historizismus, a.a.O., S. 35.

2 Vgl. derselbe, Logik der Forschung, a.a.O., S. 32. Popper begründet die Verwendung des Ausdrucks "kausale Erklärung" damit, daß die verwendeten Gesetze "die Form von Nahwirkungsgesetzen (Differentialgleichungen) haben".

3 Vgl. Ackoff, R.L., Scientific Method, a.a.O., S. 393f.

trachtenden geläufigen makroskopischen Verhalten ergänzen.¹ Was bleibt, sind Gesetzmäßigkeiten, die mit statistischen Methoden in Hinblick auf Ursache-Wirkungs-Zusammenhänge untersucht werden können.²

Die oft mit dem Kausalitätsprinzip verwechselte Erfahrung "gleiche Ursache - gleiche Wirkung"³ hat mit den hier erörterten Problemkreisen nichts gemein. Wenn darüber hinaus bedacht wird, daß beispielsweise zwei Zustände, die unter die Kennzeichnung "gleiche Ursache" fallen, strenggenommen nur dem Augenschein nach gleich sind, so erkennt man die Bedeutungslosigkeit einer solchen Regel. Jede makroskopische Betrachtung im Marketing liefert nur Mittelwerte von Ereignissen, während die Einzelheiten unbestimmt bleiben. Da aber bei jeder makroskopischen Betrachtung die Anzahl der in die Analyse einbezogenen Variablen und Parameter groß ist, resultiert daraus aufgrund des Gesetzes der großen Zahl eine geringe Schwankungsbreite, so daß die Genauigkeit eines festgestellten Quasi-Gesetzes hinreichend ist

1 Stachowiak, H., Denken und Erkennen im kybernetischen Modell, a.a.O., S. 145 ist der Auffassung, daß die Faktorenanalyse zu einer Verfeinerung des Ursachenbegriffs geführt hat, wenngleich diese Methoden vom optimalen Verhältnis zwischen Vorhersagegenauigkeit und Rethhaltigkeit noch weit entfernt sind. Weitaus pessimistischer ist Überla, der feststellt, daß nur bei Einfachstruktur der als ursächlich erklärten Variablen, die vereinfachten Beschreibungsdimensionen einer Faktorenanalyse mit den ursächlichen Einflüssen übereinstimmen können. Weiter stellt er fest: "Den Bedingungen der Koinzidenz zwischen deskriptiven, aus der Faktorenanalyse gewonnenen Dimensionen und den tatsächlichen kausalen Einflüssen wäre in weiteren Versuchen und Überlegungen nachzugehen." Überla, K., Faktorenanalyse, a.a.O., S. 293.

2 Vgl. Popper, K.R., Das Elend des Historizismus, a.a.O., S. 104ff.

3 Der Begriff "Ursache" ist "nicht nur wegen seiner inhaltlichen Unbestimmtheit, sondern auch deshalb aus der exakten wissenschaftlichen Forschung zu eliminieren..., weil sich mit ihm die mystische Vorstellung eines außerhalb des tatsächlichen Geschehens existierendes, dieses Geschehen aber zustande bringendes Agens verbindet." Stachowiak, H., Denken und Erkennen im kybernetischen Modell, a.a.O., S. 143.

und die Forderung von LITTLE nach Modellstrukturen, die die Modellergebnisse auf eine Klasse von bedeutsamen Aussagen beschränken, erfüllt sein dürfte.

3.3 Adaptivität und Kommunikation

Das von LITTLE¹ aufgestellte Postulat der Adaptivität eines Marketing-Modells besagt, daß dieses Modell vom Benutzer an neue Inputdaten angepaßt werden kann. Dabei kann es sich sowohl um Ausprägungen von Variablen als auch von Parametern handeln. Diese modellorientierte Begriffsexplikation des Terminus "Daten" besagt, daß ein adaptives Modell in der Lage sein muß, Veränderungen bei den Merkmalsausprägungen von Variablen und Parametern zu berücksichtigen. Je größer das in einem Modell zu berücksichtigende Datenvolumen ist, das sachlogisch richtig verarbeitet werden kann, um so größer ist die Adaptivität dieses Modells. Bei einem von einem Marketing-Modell zu verarbeitenden hohen Datenvolumen erscheint es zumindestens möglich, eine Ersetzung von bestimmten Daten durch andere vorzunehmen. In der Regel ist der Modelladaptivität aber dort eine Grenze gesetzt, wo sich Daten nur unter hohen Kosten beschaffen lassen, so daß die Wirtschaftlichkeit des Modelleinsatzes - verstanden als angemessenes Verhältnis zwischen Kosten der Datenbeschaffung und dem monetär bewerteten Informationsoutput des Modells - nicht mehr gegeben ist. Darüber hinaus ist das Postulat der Adaptivität eines Marketing-Modells dann gefährdet, wenn die Adaptivität in Widerspruch zur Einfachheit, Berechenbarkeit etc. steht.

Die Aufgabe der Anpassung eines Marketing-Modells an veränderte Parameterkonstellationen wird besonders von NAYLOR² betont. Er stellt dazu fest, daß allein für die Schätzung von in Marketing-Modellen aufzunehmende Parameter mindestens sieben bis zehn Beobachtungen notwendig sind, wenn eine Schätzung mit der erforderlichen Genauigkeit vorgenommen werden soll. Problem dabei ist, ob die

¹ Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

² Vgl. Naylor, T.H., Marketing-Politik-Modelle, a.a.O., S. 15f.

Schätzungen bereits auf vorhandenen, in der Datenbank des Marketing-Informationssystems gespeicherten Daten basieren, oder ob Parameterschätzungen aufgrund neu zu beschaffender Daten vorgenommen werden müssen.

Bei Modellen, die nur aus einer Gleichung bestehen, ist die Methode der kleinsten Quadrate ein geeignetes Verfahren, um die Modellparameter mit ausreichender Genauigkeit zu schätzen. In diesem Zusammenhang gewinnt das Problem der Schätzungsgüte besondere Bedeutung, da niemals auszu-schließen ist, daß bereits geringfügige Fehler bei der Parameterschätzung zu einer Verfälschung des Modelloutputs führen können. Es bedarf somit einer Vorgehensweise, die gewährleistet, daß die Güte der Schätzung überprüft werden kann.

Eine klassische Möglichkeit, die Güte von Parameterschätzungen zu überprüfen, besteht in der Anwendung von statistischen Verfahren, die Auskunft über den relativen Schätzfehler geben und ggf. Signifikanzniveaus für die Schätzung spezifizieren. Eine im Zusammenhang mit dem Einsatz von elektronischen Datenverarbeitungsanlagen bessere Vorgehensweise ist die Simulation des Modells mit alternativen Parameterkonstellationen. "Durch die Aufnahme verschiedener Ausgangswerte für die endogenen Variablen, durch Variation der Parameterwerte und/oder durch die Einführung spezifischer Häufigkeitsverteilungen für die endogenen Variablen kann das vorhandene Modell im streng kontrollierten Experiment auf neue hypothetische Situationen umgestellt werden."¹ Die Simulationsergebnisse können dann ex-post mit den tatsächlich eingetretenen Ergebnissen der Realität verglichen werden, so daß festgestellt werden kann, welche Parameterkonstellationen momentan am besten geeignet sind.

¹ Müller, W., Technik und Leistungsfähigkeit betriebswirtschaftlicher Simulationsstudien, in: Zeitschrift für Betriebswirtschaft, 38. Jg., 1968, S. 605ff, hier S. 614.

Erfordert ein anderer Untersuchungszweck eine Umformung der Modellstruktur, so können Variablen und Parameter ausgetauscht sowie Relationen und Entscheidungsregeln neu formuliert werden. Entstehende zusätzliche Zeitreihen lassen sich mit den ursprünglichen vergleichen und erlauben Aussagen über mögliche Verhaltensreaktionen des simulierten Realsystems auf veränderte Bedingungen.¹ Die Simulationsläufe können somit anzeigen, ob eine bestimmte Variable zu Recht oder zu Unrecht in das Modell aufgenommen wurde. Es kann hierbei so vorgegangen werden, daß bestimmte Variablen aus dem Modell eliminiert werden und nach dem Einfluß gefragt wird, den diese Eliminierung auf das Modellergebnis besitzt. Ist das Ergebnis negativ, kann die betreffende Variable vernachlässigt werden.

In dieser experimentellen Anwendung der Simulation wird künftig ihre große Bedeutung liegen. Das komplizierte Zusammenwirken der verschiedenen Einflußgrößen innerhalb der Unternehmung und ihrer Märkte ist für den Marketing-Manager nicht mehr überschaubar, so daß er bei seinen Schätzungen immer nur eine relativ kleine Anzahl von Faktoren beachten kann. Durch die große verarbeitungsspezifische Kapazität von Simulationsmodellen wird das Schätzproblem dadurch verbessert, daß alle als relevant angesehenen Faktoren und die Konsequenzen in den Interaktionen bei der Entscheidungsfindung berücksichtigt werden können. Die Simulation erlaubt also, durch systematische Variationen des bestehenden, im Simulationsmodell abgebildeten Systemzustandes, die Auswirkungen von unterschiedlichen Variablen und Parametern aufzuzeigen.

Unter dem Gesichtspunkt der Benutzerorientiertheit ist das Postulat der Adaptivität von Marketing-Modellen eines der bedeutsamsten benutzerorientierten Anforderungen. Es sorgt

¹ Müller, W., Technik und Leistungsfähigkeit betriebswirtschaftlicher Simulationsstudien, a.a.O., S. 614.

einerseits dafür, daß der Benutzer die Möglichkeit erhält, seinen zeitlich varianten Informationsbedarf durch Eingabe von zeitlich varianten Variablen und Parametern in das Marketing-Modell zu befriedigen. Andererseits ist durch Adaptivität der Modellansätze gewährleistet, daß ein einmal entwickeltes Marketing-Modell nicht zu schnell unanwendbar wird, nur weil es auf bestimmte Variablen und Parameter starr fixiert ist. Die Forderung nach geeigneten adaptiven Modellstrukturen impliziert aber zumindest, daß diesbezügliche Modelle im Zusammenhang mit den Testverfahren der Methodenbank gesehen werden müssen, die dazu geeignet sind, die aufzunehmenden Variablen und Parameter auf ihre Signifikanz hinsichtlich des Modelloutputs zu überprüfen.

Auf das von LITTLE¹ aufgestellte Postulat der Kommunikation des Benutzers mit den Modellen braucht in diesem Zusammenhang nicht eingegangen zu werden. Der Output von Marketing-Modellen wird in der Datenbank gespeichert, so daß die dort unter dem Gesichtspunkt der Kommunikation zwischen Datenbank und Benutzer festgestellten Sachverhalte und Forderungen hier analog zutreffen.

¹ Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470f.

3.4 Berechenbarkeit und Vollständigkeit

Modelle sind als Gleichungssysteme für die Aufnahme in die Modellbank eines Marketing-Informationssystems um so eher geeignet, wenn sie lösbar, d.h. berechenbar sind. Berechenbarkeit soll hier nicht als subjektive Fähigkeit des Marketing-Managers verstanden werden, bestimmte Rechenoperationen vorzunehmen, sondern als eine objektive Eigenschaft von Gleichungen und Gleichungssystemen. Was eine Gleichung berechenbar macht, ist die Existenz eines Berechnungsverfahrens.¹ Es ist einleuchtend, daß das Problem der Berechenbarkeit eng mit dem von LITTLE genannten Kriterium der Vollständigkeit zusammenhängt. Dieses Kriterium besagt, daß Modellstrukturen gefunden werden müssen, die es ermöglichen, viele Marketing-Phänomene zu berücksichtigen, ohne daß das Modell diese Vielfalt nicht mehr verarbeiten kann.² Diese Auffassung von LITTLE läßt sich dahingehend interpretieren, daß trotz vieler Variablen und Parameter eine Berechenbarkeit des Modells gewährleistet sein muß. Der Problemkreis der Berechenbarkeit hängt mit einem zur Verfügung stehenden Algorithmus zusammen. Aus diesem Grund ist es notwendig, den Begriff des Algorithmus näher zu präzisieren.

Es lassen sich zwei Typen von Algorithmen unterscheiden, und zwar der numerische und der logische Algorithmus.³ Numerische Algorithmen sind der Inbegriff sämtlicher Formeln zur Lösung numerischer Probleme, wobei die Art der zu vollziehenden Operationen als auch deren Reihenfolge von vornherein eindeutig festgelegt sind. Im Gegensatz dazu wird bei einem logischen Algorithmus ein zu lösendes Problem in eine Reihe von Subproblemen zerlegt, so daß man

1 Vgl. Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S. 32.

2 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-470.

3 Vgl. Aiserman, M.A. et al., Logik - Automaten - Algorithmen, München - Wien 1967, S. 312.

nach endlich vielen Lösungsschritten zu einer Lösung gelangt. Dieser Lösungsprozeß kann nach erfolgter Lösung ohne Schleifenbildung vorgenommen werden, während bei einem Anhalten des Prozesses vor Erreichen der Lösung diese selbst nicht erreicht werden kann.¹

Dieser Sachverhalt kann durch ein von NEWELL / SHAW / SIMON² entnommenes Beispiel näher verdeutlicht werden. Das Problem besteht im Öffnen eines Tresors, der ein Schloß mit 10 Sicherungen hat. Jede Sicherung wird dabei durch 100 Zahlenkombinationen bedient, so daß zur Öffnung des Panzerschranks 10^{20} mögliche Zahlenkombinationen bestehen. Der Tresor soll die verschleißbedingte Eigenschaft besitzen, daß bei einer richtig eingestellten Zahl einer Sicherung ein leichtes "Klicken" zu vernehmen ist, das das Überwinden der Sicherung anzeigt. Für jede zu überwindende Sicherung bestehen jetzt noch 100 Kombinationsmöglichkeiten, so daß der Schloßmechanismus unabhängig von den Interdependenzen der Sicherungen überwunden werden kann.

Dieses Beispiel verdeutlicht recht anschaulich, daß eine Problemlösung nach dem "Prinzip der kleinen Schritte" nur dann möglich ist, wenn die einzelnen Subprobleme voneinander unabhängig, in sich sinnvoll sind und die Anzahl dieser Subprobleme so gewählt werden kann, daß die Anzahl verbleibender Aktionen zur Erreichung des nächsten Subziels endlich ist. D.h. mit anderen Worten, daß an die Vollständigkeit des zu lösenden Problems keine unrealistischen Anforderungen gestellt werden dürfen. Darüber hinaus muß eine Bewertung der zu vollziehenden Suchstrategien möglich sein.

¹ Vgl. dazu das "Theseus-Beispiel" bei Aisermann, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 312ff.

² Vgl. Newell, A. / Shaw, J.C. / Simon, H.A., The Process of Creative Thinking, in: Gruber, H.E. / Terrell, G. / Wertheimer, M., Hrsg., Contemporary Approaches to Creative Thinking, New York 1962, S. 63ff, hier S. 76f.

Im vorliegenden Fall ist der Tresor nur dann zu öffnen, wenn die Zahlen der jeweiligen Sicherungen notiert werden, die einmal ausprobiert worden sind. In der skizzierten Vorgehensweise, den Geldschrank zu öffnen, steckt allerdings eine gewisse Willkür. Es ist durchaus möglich, beim Öffnen mit unterschiedlichen Sicherungen zu beginnen, und zwar danach, wie sich diese durch Geräusche beim Überwinden der Sicherung "verraten". Im Gegensatz zu sonstigen numerischen Problemen bleibt es der Willkür des einzelnen überlassen, mit welcher Sicherung er beginnen will. Verschiedene Personen können somit durchaus auf unterschiedliche Art und Weise die Öffnung des Geldschrankes erreichen. Aus diesem Beispiel ergeben sich eine Reihe von Eigenschaften, über die ein Algorithmus verfügen muß.

"Von einem Algorithmus wird verlangt, daß das durch ihn beschriebene Lösungsverfahren so exakt und allgemein verständlich ist, daß keinerlei Willkür übrigbleibt."¹ Das Verfahren selbst muß intersubjektiv in Form endlich vieler Regeln mitteilbar sein, so daß für jeden zu vollziehenden Verfahrensschritt eindeutige Anweisungen vorliegen. Da Willkürfreiheit vorausgesetzt wird, bilden die vollzogenen Lösungsschritte einen "determinierten Prozeß"², der jederzeit von Dritten mit demselben Ergebnis vorgenommen werden kann. Weiterhin dient ein Algorithmus nicht nur der Lösung eines einzelnen Problems, sondern der Lösung einer ganzen Klasse von identischen Problemen. Die Verfahrensregeln werden dabei auf Anfangswerte angewendet, die von Fall zu Fall variieren können. Das Auffinden dieser pro-

1 Aiserman, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 315.

2 Die folgenden Ausführungen lehnen sich an dieselben, Logik - Automaten - Algorithmen, a.a.O., S. 315f. an.

blemklassen-spezifischen Algorithmen ist das eigentliche Problem.¹

Darüber hinaus wird von einem Algorithmus verlangt, daß er nach endlich vielen Schritten anhält. Aus dieser Forderung nach dem Abbrechen eines Algorithmus folgt der Begriff des Anwendungsbereiches dieses Algorithmus.² Unter Anwendungsbereich soll der größte Bereich derjenigen Anfangswerte verstanden werden, für die der Algorithmus abbricht. Wenn die Bedingungen des betreffenden Problems aus dem Anwendungsbereich entnommen sind, so verarbeitet der Algorithmus diese Bedingungen in der Lösung des Problems und hält an. Dieses wird als Lösungsgarantie des Algorithmus bezeichnet. Stammen dagegen die Bedingungen nicht aus dem Anwendungsbereich, so gelangt der Algorithmus zu keinem Ende und liefert somit keine Lösung. Es ist zu beachten, daß die Anzahl der bei einem Algorithmus zu vollziehenden Operationen a-priori unbekannt ist und daher nur als potentiell möglicher Prozeß zu verstehen ist.

Aus dieser Perspektive wird deutlich, daß das von LITTLE aufgestellte Postulat der Vollständigkeit zwar aus der Sicht der materiellen Marketing-Problemstellung wünschenswert ist, seine Realisierung in Gleichungssystemen von Marketing-Modellen aber auf Schwierigkeiten stoßen kann. Dieses gilt insbesondere dann, wenn die für eine Lösung aufzuwendende Zeit und die rechentechnischen Schwierigkeiten überproportional anwachsen sollten.

¹ "Das generell zugrundeliegende Problem kann folgendermaßen formuliert werden: 'Existiert ein Algorithmus, der für eine beliebige Formel eines logischen [und numerischen] Kalküls entscheidet, ob sie aus vorgegebenen Axiomen folgt oder nicht?'. " Niedereichholz, J., Betriebswirtschaftslehre und Informatik, in: Wissenschaftsprogramm und Ausbildungsziele der Betriebswirtschaftslehre, a.a.O., S. 225ff, hier S. 225.

² Vgl. Aiserman, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 317.

Diese drei Eigenschaften eines Algorithmus, die Lösungsgarantie¹, die Allgemeingültigkeit und die Endlichkeit sind Eigenschaften, die man bei allen untersuchten Algorithmen festgestellt hat. Diese Berechenbarkeitspostulate lassen sich anschaulich am Beispiel einer Turing-Maschine erläutern. Eine n -stellige Funktion heißt dann Turing-berechenbar², wenn es dieser Maschine nach endlich vielen Schritten gelingt, eine Rechnung vorzunehmen, d.h. anzuhalten. Turing-Berechenbarkeit besagt, daß diese Funktion dann berechenbar ist, wenn ein Algorithmus, der sich auf die Berechnung einer ganzzahligen Funktion auf einer Turing-Maschine zurückführen läßt³, existiert.

Im folgenden soll der Begriff der Berechenbarkeit auf eine andere Art expliziert werden⁴, um zu untersuchen, inwieweit das Postulat der Vollständigkeit von LITTLE unter dem Primat der Berechenbarkeit berücksichtigt werden kann.

Ziel ist es, möglichst umfangreiche Klassen von Funktionen aufzustellen, die sich durch das Kriterium der Berechenbarkeit auszeichnen. Gelingt es nicht, die durch fixierte und nur für gewisse Operationen zulässigen Ausgangsfunktionen zu erweitern, so hat man eine Explikation des Berechenbarkeitsbegriffs erhalten. Diese Operationen führen wiederum zu berechenbaren Funktionen, die man ausgehend von Ausgangsfunktionen durch einmaliges Anwenden der Operationen erhält (Rekursivität). Diese

1 Vgl. dazu Reitman, W.R., Heuristic Decision Procedures, Open Constraints, and the Structure of Ill-Defined Problems, in: Shelly, M.W. / Glenn, L.B., Hrsg., Human Judgements and Optimality, New York - London 1964, S. 282ff, hier S. 303.

2 Vgl. Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S. 36.

3 Vgl. dazu Hermes, H., Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit, a.a.O., S. 76ff.

4 Vgl. dazu die umfassenden Darstellungen von Gödel, K. / Church, A. / Kleene, S.C. und Rosner, B. bei Stegmüller, W., Hrsg., Unvollständigkeit und Unentscheidbarkeit, Wien 1959.

Explikation des Berechenbarkeitsbegriffs von Funktionen durch Rekursivität kann mittels der CHURCHschen These¹ darauf zurückgeführt werden, daß sämtliche Explikationen des "Berechenbarkeitsbegriffs"- obgleich unterschiedlich definiert - die gleichen Extensionen haben. Das mathematische Problem der Berechenbarkeit ist dann gelöst, wenn es einen Satz gibt, auf den sämtliche Berechenbarkeitsexplikate zurückgeführt werden können.² Die Explikation des Berechenbarkeitsbegriffs hat somit einige Entwicklungsstufen durchlaufen, und zwar die der Turing-Berechenbarkeit, die der elementaren und primitiv-rekursiven Funktionen und die der allgemein-rekursiven Funktionen, bis durch die CHURCHsche These der Beweis gelang, daß jede effektiv berechenbare Funktion allgemein-rekursiv ist. Daraus folgt, daß die von verschiedenen Autoren vorgeschlagenen Präzisierungen des Algorithmusbegriffs einander äquivalent sind.³

Somit läßt sich folgende Bedingung formulieren, wenn Berechenbarkeit einer Funktion gegeben sein soll:

Eine Funktion $f(x_1; x_2; \dots; x_n)$ heißt allgemein-rekursiv, wenn ein endliches System von Gleichungen existiert, so daß sich für jeden Satz von Zahlen $x_1^+; x_2^+; \dots; x_n^+$ eine

1 "Jede effektiv berechenbare Funktion (jedes effektiv entscheidbare Prädikat) ist allgemeinrekursiv." Kleene, S. C., Introduction to Metamathematics, 2. Aufl., Amsterdam 1964, S. 300. Vgl. dazu auch die Interpretation bei Aiserman: "Wegen Äquivalenz der Begriffe 'rekursive Funktionen' und 'auf einer Turing-Maschine berechenbare Funktionen' kann man nach der CHURCHschen These den Begriff Algorithmus auch folgendermaßen definieren: Ein Algorithmus ist jede Prozedur, die sich auf die Berechnung einer ganzzahligen Funktion auf einer entsprechenden Turing-Maschine zurückführen läßt." Aiserman, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 386. (Der letzte Satz des Zitates im Original kursiv).

2 Vgl. Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S. 31 und 43.

3 Vgl. Hermes, H., Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit, a.a.O., S. 76.

und nur eine Zahl y^+ finden läßt, so daß die Gleichung $y = f(x_1^+; x_2^+; \dots; x_n^+)$ durch endlich viele Anwendungen von Operation 1 (Einsetzen von Zahlen an Stelle von Veränderlichen) und Operation 2 (Zulassen beliebiger Rekursions-schemata zur Erzeugung von Funktionen aus den bereits gewonnenen) aus dem Gleichungssystem ableitbar ist.¹

In dieser Definition wird verlangt, daß das Gleichungssystem einen gegebenen Funktionswert mit Hilfe anderer Funktionswerte und den Werten von Hilfsfunktionen so bestimmt, daß der Funktionswert an allen Punkten aufgrund dieses Systems eindeutig berechnet werden kann.²

Diese Ausführungen zeigen, daß beim Einsatz von Modellen in der Modellbank eines Marketing-Informationssystems Berechenbarkeitsprobleme von Funktionssystemen nicht vernachlässigt werden dürfen. Das von LITTLE aufgestellte Postulat der Vollständigkeit ist nur dann ein sinnvolles Kriterium, wenn die Modelle nicht zu komplex werden, daß entweder keine Lösung zu erzielen oder die Lösungen mit zu großem Zeitaufwand verbunden sind. Die in mathematischen Marketing-Modellen auftretenden Probleme sind fast immer Wahlprobleme, bei denen sich der Entscheidungsträger zwischen mehreren unterschiedlichen Zuständen (Turing-Berechenbarkeit!) zu entscheiden hat. Er wird versuchen, ein Optimum oder zumindest eine "gute" Lösung zu erreichen, die gemessen an der vorgegebenen Zielsetzung, dem Optimum möglichst nahe kommt. Dabei wird vorausgesetzt, daß die zur Vornahme dieser rationalen Wahlhandlung benötigten Daten immer zur Verfügung stehen, so daß ein auf völliges Unwissen beruhendes "blindes Zufallshandeln" ausgeschlossen ist. Aufgrund der Allgemeinrekursivität hängt die Verwendung eines bestimmten Verfahrens zur Optimumbestimmung

1 Vgl. Aisermann, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 357.

2 Vgl. dazu das Beispiel bei Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S. 357.

von der Genauigkeit der zur Verfügung stehenden Daten - des Satzes von Zahlen - und der Anzahl der zu vollziehenden Rechenoperationen ab, um zu einer Lösung zu gelangen. Bei Kenntnis der zukünftigen Datenentwicklung besteht zwischen dem Zielfunktionswert und dem Mitteleinsatz eine deterministische Beziehung, so daß sich aufgrund des Prinzips der Allgemeinrekursivität das Gleichungssystem lösen lassen muß.

Besteht hingegen über die zukünftige Datenentwicklung Unsicherheit, ist keine funktionale Beziehung mehr zwischen dem Zielfunktionswert und dem Mitteleinsatz gegeben. In diesem Fall ist es nicht mehr durch Anwendung von Rekursionsschemata möglich, zu einer Lösung zu gelangen. Die Richtigkeit dieser Aussage läßt sich durch eine plausible Mächtigkeitsüberlegung zeigen.¹ Da es in diesem Fall sicherlich nur abzählbar viele Verhaltensfunktionen und somit auch nur abzählbar viele rekursive Funktionen gibt - im Gegensatz dazu aber überabzählbar viele Funktionen zahlen-theoretischer Art bekannt sind - muß es auch überabzählbar viele nichtberechenbare Funktionen geben.² Durch diesen Existenzbeweis ist gezeigt worden, daß es bei Unsicherheit über die zukünftige Datenentwicklung keine mathematischen Modellösungen geben kann. Erst durch die Tatsache, daß man innerhalb bestimmter Grenzen den Eintritt eines Datums für möglich erachtet, sind analytische Lösungsverfahren möglich. Es handelt sich aber hier um abzählbar viele Berechenbarkeitsmöglichkeiten. Deshalb bietet sich aus betriebswirtschaftlichen Gründen diese Lösung nicht an - es existieren zu viele Ziel-Mittel-Kombinationen, die die Wirtschaftlichkeit der Lösung in Frage stellen. Zur Lösung eignen sich aus diesem Grund nur auf den konkreten Fall zugeschnittene Lösungsalgorithmen, die bei bestimmten

1 Vgl. Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S. 47.

2 Vgl. ebenda, S. 47f.

Problemstellungen eine "befriedigende" Lösung gewährleisten.¹

Es kann somit festgestellt werden, daß bestimmte Rekursionsschemata, die zwangsläufig bei nicht-programmierbaren Entscheidungen zu Lösungen führen, nicht vorhanden sind. Bei derartigen Problemkreisen steht die individuelle Gestaltung des Lösungsweges im Vordergrund, wobei allgemeine Regeln in Form rekursiver Algorithmen a-priori unbekannt sind. Bei diesen mehr oder weniger innovativen Entscheidungen verfügt der Entscheidungsträger über kein Verfahren, das es ihm ermöglicht, aus einer Vielzahl möglicher Alternativen eine bestimmte auszuwählen.

Im Zusammenhang mit der Frage der Berechenbarkeit soll hier noch einmal auf den unpräzisen Begriff der heuristischen Verfahren² eingegangen werden, die sich dadurch auszeichnen, daß darunter sämtliche Prinzipien und Vorgehensweisen subsumiert werden, die zu einer im Durchschnitt reduzierten Suche nach "befriedigenden" oder geeigneten Lösungen führen, ohne daß die Gewähr dafür besteht, daß eine derart charakterisierte Lösung gefunden wird.³ Bei diesen Verfahren spielen die Intuition und die Erfahrung des Entscheidungsträgers eine nicht zu unterschätzende Rolle. Insbesondere gilt es zu klären, ob das Kriterium eines a-priori festlegbaren Algorithmus hinreichend ist, um die vorgenommene Klassifizierung zwischen analytischen, quasi-analytischen und heuristischen Verfahren zu rechtfertigen.

1 Vgl. dazu z.B. Cloonan, J.B., A Heuristic Approach to Some Sales Territory Problems, in: Proceedings Of The Fourth International Conference On Operational Research, hrsg. v. Hertz, D.B. / Malese, J., New York usw. 1966, S. 284ff, hier insbesondere S. 285ff und Kuehn, A.A. / Hamburger, M.J., A heuristic program for locating warehouses, in: Management Science, Vol. IX, 1962/63, S. B-643ff.

2 Vgl. dazu den Punkt 11 dieses Abschnittes

3 Vgl. Klein, H.K., Heuristische Entscheidungsmodelle - Neue Techniken des Programmierens und Entscheidens für das Management, Bd. 9 der Schriftenreihe "Die Betriebswirtschaft in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1971, S. 35f.

Als finales Abgrenzungskriterium zwischen diesen Verfahren kann die Existenz eines vor Lösungsbeginn festlegbaren Algorithmus angesehen werden. Bei analytischen Verfahren ist dies für das gesamte Problem, bei quasi-analytischen nur bei Unterproblemen möglich, während bei heuristischen Verfahren ein Lösungsalgorithmus fehlt. Eine derartige finale Abgrenzung führt aber zu dem Problem, daß als Unterscheidungskriterium ein Begriff gewählt wird, der selbst mit Unschärfe behaftet ist. So ist festzustellen, daß es bis heute nicht gelungen ist, eine mathematische - d.h. auf eine Fülle von Assoziationen verzichtende Beschreibung des Algorithmusbegriffs zu geben. Gängige Definitionen wie: "Ein Algorithmus liegt genau dann vor, wenn gegebene Größen, auch Eingabegrößen, Eingabeinformationen oder Aufgaben genannt, aufgrund eines Systems von Regeln, Umformungsregeln, in andere Größen, auch Ausgabegrößen, Ausgabeinformationen oder Lösungen genannt, umgeformt oder umgearbeitet werden können"¹ reichen zwar annähernd aus, um zu entscheiden, ob ein vorliegendes Verfahren ein Algorithmus ist oder nicht. Sie lassen aber noch bestimmte Freiheiten, um Verfahren zum Algorithmus zu erklären oder nicht. Aus diesem Grund muß die Begriffsexplikation von analytischen, quasi-analytischen und heuristischen Verfahren mit derselben Unschärfe behaftet sein wie eine verbale Algorithmusdefinition, die nur mathematisch exakt zu explizieren ist.

Es ist bekannt, daß sich viele Gleichungen durch Näherungsverfahren lösen lassen. Bei einer derartigen Näherung geht man von einer vorläufigen Lösung aus, die dann durch das Näherungsverfahren selbst in eine genaue bzw. genauere Lösung transformiert wird. Diese Methode der sukzessiven Approximationen kann als heuristisches Lösungsprinzip bezeichnet werden.

¹ Art.: Algorithmus, in: Kleine Enzyklopädie der Mathematik, Leipzig 1965, S. 779, zitiert bei Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 261.

Auch das quasi-analytische Verfahren ist nichts anderes als die schrittweise Näherung eines Problems. Es wird so vorgegangen, als ob eine Lösung mittels eines Algorithmus erzielbar wäre, wobei die Problemstruktur als so gelagert angesehen wird, daß unmittelbar durch einen rekursiven Algorithmus eine Quasi-Lösung erreicht wird. Angesichts dieser terminologischen Unklarheiten wird vorgeschlagen, nur noch zwischen analytischen und approximativen Verfahren zu unterscheiden.

Darüber hinaus ist insbesondere zur Abgrenzung zwischen quasi-analytischen und heuristischen Verfahren (den approximativen Verfahren) die getroffene Unterscheidung zwischen numerischen und logischen Algorithmen von besonderer Bedeutung.¹ GÖDEL² hat den Beweis erbracht, daß jeder logische Algorithmus mittels eines von ihm entwickelten Verfahrens auf einen numerischen Algorithmus zurückgeführt werden kann. "Bei der Vervollkommnung dieser Methoden stellte sich dann heraus, daß sich überhaupt jeder Algorithmus auf einen numerischen Algorithmus zurückführen läßt."³

Dieses als "Gödelisierung" bezeichnete Verfahren ist als eine Erweiterung von Berechenbarkeits-Überlegungen auf nicht numerische Bereiche zu verstehen.⁴ Ziel der Gödelisierung ist es, jeden effektiv gegebenen Bereich durch eine effektiv berechenbare Funktion abzubilden, wobei diese Abbildung eineindeutig durch natürliche Zahlen (Gödelnummern) vorgenommen wird. Nach Einführung der Gödelnummern hat man es nur noch mit natürlichen Zahlen zu tun und nicht mehr mit den oft unterschiedlichen Schreibweisen

1 Vgl. Aiserman, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 312.

2 Vgl. dazu Gödel, K., Über formal unentscheidbare Sätze der Principia Mathematica und verwandte Systeme, in: Monatshefte für Mathematik und Physik, 38. Jg., 1931, S. 173ff.

3 Aiserman, M.A. et al., Logik - Automaten - Algorithmen, a.a.O., S. 329.

4 Vgl. Behnke, H. et al., Mathematik, Bd. 2, a.a.O., S.46.

einzelner Bedingungen und Lösungen.¹ Es kann in diesem Fall von einem Algorithmus gesprochen werden, der die Nummern von formulierten Bedingungen in die Nummer einer formulierten Lösung transformiert, wobei es sich um einen numerischen Algorithmus handelt.

Die Anwendbarkeit der Gödelisierung erstreckt sich aber nicht nur auf die Behandlung numerischer Probleme; sondern auch nicht numerische Phänomene - wie die Sprachen der mathematischen Logik - sind im Sinne des Verfahrens effektiv gegebene Bereiche. Die "Algorithmen" dieser Sprache sind letztlich nur Syntaxregeln, die mittels eines normierten Alphabets aufgebaut werden können.

Aus diesem Gesichtspunkt heraus, daß jeder logische Algorithmus durch Gödelisierung auf einen numerischen Algorithmus zurückgeführt werden kann, ist eine Unterscheidung zwischen analytischen und heuristischen Verfahren auf Basis eines Algorithmus vorzunehmen. Die analytischen und approximativen Verfahren unterscheiden sich nicht dadurch, daß einerseits ein numerischer Algorithmus vorhanden, andererseits keiner vorhanden ist, sondern allein durch die Tatsache, daß dieser Algorithmus bei analytischen Verfahren vorgegeben ist, während er bei approximativen Verfahren im Laufe der Lösung als numerischer bzw. logischer (quasi-analytisches Verfahren) oder als logischer (heuristisches Verfahren) aufgebaut wird.

Dieses ist ein Grund dafür, daß approximative Verfahren versagen müssen, wenn sie ein lokales Optimum ansteuern sollen.² Darüber hinaus ergeben sich bei Existenz mehrerer Variablen hinsichtlich der multidimensionalen Ober-

1 Vgl. zu den folgenden Ausführungen Behnke, H. et al., a.a.O., S. 46ff.

2 Vgl. Minsky, M., Steps Toward Artificial Intelligence, in: Feigenbaum, E.A. / Feldman, J., Hrsg., Computers and Thought, New York 1963, S. 406ff, hier S. 411.

fläche weitere Probleme bei der Optimumbestimmung.¹ Diese Verfahren geraten trotz "Arithmetisierbarkeit" äußerst schnell an die Grenze ihrer Anwendungsmöglichkeiten, da der sich damit verbundene Rechenaufwand exponentiell erhöht.

Dem Gedanken, daß in bestimmten Situationen weder Optimierungs- noch approximative Lösungsverfahren anwendbar sind, versucht die Simulation Rechnung zu tragen. Simulationsverfahren zielen darauf ab, mittels a-priori bekannten Algorithmen, Hypothesen über das Verhalten von Systemen zu generieren. So hat beispielsweise FORRESTER² untersucht, wie sich das Marketing-System bei Veränderung bestimmter Parameter der Lagerhaltung, des Umsatzes etc. verhält. Eine Problemlösung ist mittels Simulationstechniken nur dann zu erreichen, wenn Simulations- und approximative Verfahren zusammen angewendet werden. In diesem Fall haben die zur Anwendung gelangenden approximativen Verfahren - die überdies stark von der individuellen Problemstruktur des Einzelfalls abhängen - die Aufgabe, den Problemlösungsprozeß zu unterstützen. Das Spezifikum der approximativen Verfahren besteht in einer derartigen Situation darin, ein schwer zu erreichendes Ziel in leichter zu erreichende Subziele zu zerlegen. Die sich dadurch manifestierende "Politik der kleinen Schritte" besteht u.a. in der Abweichungsanalyse zwischen dem gewünschten Ergebnis und dem erreichten Istzustand. Nach KIRSCH³ werden die bestehenden Abweichungen dann in kleine, leichter zu lösende Abweichungen aufgespalten. Dieses Prinzip - durch iterativ-rekursive Zwischenschritte, den Entscheidungsträger einer

1 Diesbezügliche Überlegungen finden sich bei Hooke, R. / Jeeves, T.A., Comment on Brooks' Discussion of Random Methods, in: Operations Research, Vol. 6, S.881f, hier S. 881 und Wilde, D.J., Optimum Seek Methods, Englewood Cliffs, N.J. 1964, S. 61f.

2 Vgl. Forrester, J.W., Industrial Dynamics, 2. Aufl., Cambridge, Mass. 1962, S. 24ff.

3 Vgl. Kirsch, W., Entscheidungsprozesse, Bd. II, a.a.O., S. 177.

Lösung näherzubringen - bildet die Basis für die Theorie des "Human Problem Solving"¹, die in unterschiedlich gelagerten Problemlösungsprogrammen wie dem "Logic Theorist"² oder ihrer Weiterentwicklung zum "General Problem Solver"³ (GPS) ihre Konkretisierung findet.

¹ Vgl. Newell, A. / Shaw, J.C. / Simon, H.A., The Process of Creative Thinking, in: Gruber, H.E. / Terrell, G. / Wertheimer, M., Hrsg., a.a.O., S. 63ff, insbesondere S. 77ff.

² Vgl. dieselben, Chess-Playing Programs and the Problem of Complexity, in: Feigenbaum, E.A. / Feldman, J., Computers and Thought, New York 1963, S. 39ff.

³ Ernst, G.W., GPS and Decision Making - An Overview, in: Banerji, R.B. / Mesarovic, M.D., Hrsg., Theoretical Approaches to Non-Numerical Problem Solving, Heidelberg - New York 1970, S. 59ff.

3.5 Integrationsfähigkeit

Unter Modellintegration soll hier die "Weiterverarbeitbarkeit" von End- oder Zwischenergebnissen eines Modells in nachfolgenden Modellen verstanden werden. Dieser Integrationsbegriff ist nicht auf ein Gesamtunternehmensmodell abgestellt, das sämtliche Interdependenzen mit anderen Unternehmensbereichen umfaßt, sondern betrifft ausschließlich die Voraussetzungen und Grenzen der Integration von erstens Prognosemodellen, zweitens Entscheidungsmodellen und drittens von Prognose- und Entscheidungsmodellen. Die hier zu lösende Aufgabe besteht in erster Linie in der Erforschung der Ursachen, die bislang eine verstärkte Modellintegration verhindert haben. Gelingt es, diese Ursachen zu identifizieren, so lassen sich Wege zu einer verstärkteren Verzahnung von Marketingmodellen untereinander aufzeigen.

Bei der Untersuchung von Integrationsmöglichkeiten sind zwei Problemkreise zu unterscheiden. Der erste betrifft die mehr formalen mathematischen Voraussetzungen der Modellintegration, während sich der zweite Problemkreis mehr damit befaßt, ob und inwieweit eine Modellintegration überhaupt sinnvoll ist. Beide Problemkreise sind interdependent und lassen sich letztlich nicht exakt voneinander abgrenzen. Zunächst soll der erste Problemkreis erörtert werden.

Der primäre Grund, warum sich bei der Integration von Prognosemodellen erhebliche Schwierigkeiten ergeben, hat seine Ursache in der Inkompatibilität der diesen Modellen zugrundeliegenden Konzeptionen. Im Laufe seiner Entwicklung hat sich das Marketing wiederholt mit der Integration verschiedener Theoriekonzeptionen befaßt.¹ Die geleisteten Ansätze können aber für sich genommen, keine absoluten Prio-

¹ Vgl. Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, a.a.O., S. 167.

ritätsansprüche erheben, so daß die Suche nach einem umfassenden Bezugsrahmen und nach Gestaltungsgrundsätzen auf höherer Ebene fortgeht.¹ Der als Bezugsrahmen postulierte entscheidungs- und systemorientierte Ansatz ist aber bei der Modellentwicklung nur insoweit von Bedeutung, als es gelingt, ihn konkret-rechnerisch umzusetzen. Die in neuerer Zeit erbrachten kybernetischen Modellansätze krankten an der einseitigen Anwendbarkeit ihrer Lösungsalgorithmen. Bislang ist es nur gelungen, Bestellmengen- und Lagerhaltungsprobleme mittels Operatorenrechnung konkret-rechnerisch zu lösen.² Dieses ist u.a. darauf zurückzuführen, daß sich nur wenige Marketing-Phänomene dergestalt abbilden lassen, daß die Anwendung eines spezifischen mathematischen Verfahrens, das Implikationen hinsichtlich einer bestimmten Art der Abbildung hat, möglich ist.

Voraussetzungen einer konkret-rechnerischen Erfassung von regelungstheoretisch interpretierten Marketingphänomenen wie der graphische Entwurf eines Regelkreismodells, Ermittlung der günstigsten Variante dieses Modells und Vorgabe dieser Variante als mathematisches Modell zur Steuerung und Regelung konkreter Marketingprozesse, sind bislang nur bei quasi-deterministischen Prozessen wie eben bei der Lagerhaltung und der Bestellmengenpolitik gegeben. Es gibt zwar eine Reihe von unterschiedlichen Prognosemodellen für identische Marketing-Phänomene, die sich aber durch verfahrens- und problemspezifische Prämissen voneinander unterscheiden.

1 Vgl. Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, a.a.O., S. 167.

2 Vgl. hierzu Baetge, J. / Steenken, H.-U., Theoretische Grundlagen eines Regelungsmodells zur operationalen Planung und Überwachung betriebswirtschaftlicher Prozesse, in: Zeitschrift für betriebswirtschaftliche Forschung, 23. Jg., 1971, S. 593ff; dieselben, Regelungstheoretischer Ansatz zur operationalen Planung und Überwachung von Produktion und Lagerhaltung, in: Zeitschrift für betriebswirtschaftliche Forschung, 24. Jg., 1972, S. 22ff; Schiemenz, B., Regelungstheoretische Grundkonzepte und ihre Anwendung zur Gestaltung von Produktions-Lagerhaltungssystemen, in: Zeitschrift für betriebswirtschaftliche Forschung, 24. Jg., 1972, S. 434ff.

Diese heterogenen Prämissen können einerseits die Integration von Prognosemodellen erschweren, andererseits sogar unmöglich machen. So sind die Prämissen eines marginal-analytischen Modells (Stetigkeit, Differenzierbarkeit etc.) mit denen eines Modells der linearen Programmierung (Ganzzahligkeit, Nichtnegativität etc.) wenig kompatibel. Somit läßt sich folgende These aufstellen:

Die Integration von Prognosemodellen setzt identische Prämissen bei den einzelnen zu integrierenden Modellen voraus (Prämissenkompatibilität).

Bei realitätsnahen Prognosemodellen handelt es sich überwiegend um solche, die sich durch mehr oder weniger schnell im Zeitablauf ändernde dynamische Relationen auszeichnen. Mittels dieser dynamischen Relationen gelingt die teilweise Berücksichtigung sich ständig ändernder Gegebenheiten. Da Prognosemodelle aber nur Daten aus den für das jeweilige Modell relevanten Segmenten der Umwelt verarbeiten, nehmen die Integrationsprobleme aufgrund unterschiedlicher Datenentwicklungen zu bzw. ändern sich im Zeitverlauf. Diese mangelnden Integrationsmöglichkeiten, die sich auf die für die jeweiligen Modelle benötigten Daten zurückführen lassen, können nur dann behoben werden, wenn die Modelle mit identischen Daten arbeiten oder diese ineinander umrechenbar sind. Eine Umrechenbarkeit ist immer dann gegeben, wenn Daten verschiedener Modelle auf identischen statistischen Bezugsbasen beruhen.¹ Diese Daten werden dann in Form von Variablen oder Parametern in den jeweiligen Prognosemodellen berücksichtigt. Die Analyse der in

¹ Vgl. z.B. Adam, A., Messen und Regeln in der Betriebswirtschaft - Einführung in die informationswissenschaftlichen Grundzüge der industriellen Unternehmensforschung, Würzburg 1959, S. 55. Adam stellt hier an die Indizes folgende Forderung: "Wir richten an die Indizes eine syntaktische Forderung: es soll zur Verkettung der Elementargrößen eines Systems eine beziehungsgetreue Verkettung der Indizes dieser Elementargrößen existieren"

Prognosemodellen enthaltenen Variablen führt zur folgenden These:

Die Integration von Prognosemodellen setzt voraus, daß die zu erklärende Variable eines Modells in anderen Modellen erklärende Variable ist (Variablenkompatibilität).

Diese These soll im folgenden näher erläutert werden.

Bei der Integration von Prognosemodellen ist zu berücksichtigen, daß ein bestimmtes Prognosemodell den Input anderer Prognosemodelle liefern kann, wenn sich die Modelle gegenseitig aufeinander beziehen lassen. In diesem Fall ist es möglich, eine erklärende Variable eines Modells durch die zu erklärende Variable eines anderen Modells zu ersetzen. Ein Beispiel soll dieses näher verdeutlichen.

Es wird angenommen, daß drei Prognosemodelle vorhanden sind, die sich dadurch auszeichnen, daß die zu erklärende Variable des ersten Modells im zweiten Modell erklärende Variable ist, die zu erklärenden Variablen des ersten und zweiten Modells im dritten Prognosemodell erklärende Variablen sind. Formal lassen sich diese Bedingungen wie folgt schreiben:

Prognosemodell 1: Prognosemodell 2: Prognosemodell 3:

$$A = f(x)$$

$$B = f(A, y)$$

$$C = f(A, B, z)$$

Das erste und das zweite Modell lassen sich nur dann integrieren, wenn in die Funktion des zweiten Modells für die Variable A die funktionale Beziehung des ersten Modells $f(x)$ eingesetzt wird. Das zweite Prognosemodell hat dann die Form

$$(1) \quad B = f[f(x), y] .$$

In analoger Weise lassen sich das zweite und das dritte Modell in der Form

$$(2) \quad C = f[A, f(A, y), z]$$

und das erste, zweite und dritte Modell in der Form

$$(3) \quad \begin{aligned} C &= f[f(x), f(A, y), z] \\ &= f[f(x), f(f(x), y), z] \end{aligned}$$

zusammenfassen.

Eine derartige Integration, auf deren Grenzen noch eingegangen wird, ist daran gebunden, daß die zweite These zutrifft.

Es soll nun das Problem erörtert werden, unter welchen Bedingungen eine Integration von Prognosemodellen möglich ist, wenn die zweite These nicht zutrifft. In diesem Fall lassen sich a-priori zwischen den Variablen keine Beziehungen herstellen. Auch hier soll ein Beispiel den Sachverhalt verdeutlichen. Formal lauten die Prognosemodelle:

Prognosemodell 1: Prognosemodell 2: Prognosemodell 3:

$$A = f(x)$$

$$B = f(y, z)$$

$$C = f(u, v, w)$$

Ohne die Kenntnis von Relationen zwischen den Variablen dieser drei Prognosemodelle ist eine Modellintegration unmöglich.

Wird angenommen, daß sich aufgrund ökonomischer Untersuchungen ein hinreichend straffer Zusammenhang zwischen y und A , nämlich $y = g(A)$ und zwischen v und z , nämlich $V = g(z)$ ergeben hat, so lassen sich die Prognosemodelle integrieren. Das erste und das zweite Modell lassen sich

in der Form

$$(4) \quad B = f[g(a), z]$$

und das zweite und dritte Modell in der Form

$$(5) \quad C = f[u, g(z), w]$$

zusammenfassen. Eine weitergehende Modellintegration ist nicht möglich, da keine weiteren Relationen zwischen den Modellen bekannt sind.¹ Es läßt sich somit folgende These aufstellen:

Eine Integration von Prognosemodellen setzt voraus, daß mindestens zwischen einer zu erklärenden Variablen eines Modells und einer erklärenden Variablen eines anderen Modells eine Relation vorhanden ist (Relationenkompatibilität).

Ein typisches Beispiel einer derartigen Modellintegration ist ein preispolitisches Modell, das sowohl den Absatz - als auch den Beschaffungsmarkt berücksichtigt. Über eine Produktionsfunktion $x = f(r_1, r_2, \dots, r_n)$, die als empirisch-technische Relation interpretiert werden kann, wird der Absatzmarkt über die Variable x mit dem Beschaffungsmarkt über die Faktoreinsatzmengen $r_1, r_2, \dots, r_n$ gekoppelt. Diese preispolitischen Modelle ermöglichen die Bestimmung des Gewinnmaximums, wenn die anbietende Unternehmung z.B. auf dem Beschaffungsmarkt Monopsonist, auf dem Absatzmarkt Atomist ist.

Dieses Beispiel zeigt, daß eine Modellintegration immer dann vorgenommen werden kann, wenn zwischen bestimmten Variablen eines Modells und bestimmten Variablen eines anderen Modells ein wahrscheinlich kausaler Zusammenhang an-

¹ Zum Zusammenhang zwischen dem Kausalitätsprinzip und der Relationslogik vgl. Stachowiak, H., Denken und Erkennen im kybernetischen Modell, a.a.O., S. 143 und die dort angegebene Literatur.

genommen oder getestet werden kann. Das Problem der Modellintegration läßt sich somit hinsichtlich der dritten These darauf zurückführen, daß empirische Relationen zwischen den zu integrierenden Prognosemodellen festgestellt werden können, die sich bei empirischen Tests als hinreichend straff erwiesen haben. Vage stochastische Beziehungen werden es in der Regel nicht gestatten, Prognosemodelle - insbesondere wenn diese deterministisch sind - zu integrieren.

Die für die Integration von Prognosemodellen notwendigen Voraussetzungen, von denen entweder alle drei zusammen oder die erste und die zweite oder die zweite und die dritte gegeben sein müssen, sind in folgender Abbildung zusammenfassend dargestellt.

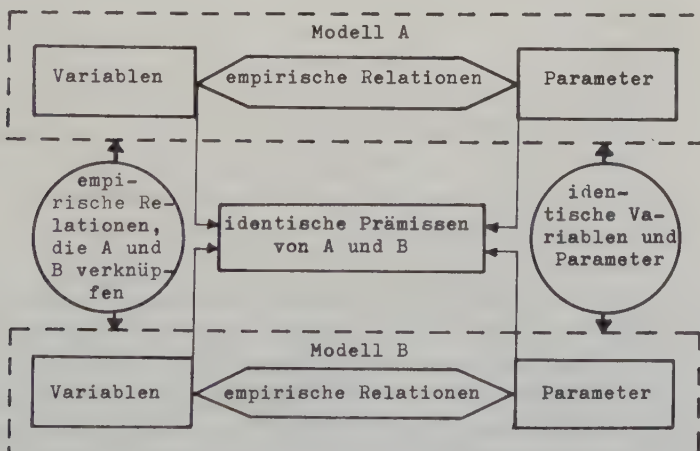


Abb.19 : Integrationsbedingungen von Prognosemodellen

Das zu lösende Problem bei der Integration von Entscheidungsmodellen besteht in erster Linie darin, ob und inwieweit die Zielfunktionen sich integrieren lassen. Die Wahl des Kriteriums für eine kardinale Ordnung zulässiger

Handlungsalternativen setzt sich aus zwei Wahlakten zusammen, und zwar erstens aus der Wahl der Zielgröße(n) und zweitens aus der Wahl der Intensität, mit der diese Zielgrößen angestrebt werden.

Vom ersten Wahlakt soll zunächst angenommen werden, daß er bereits bei den jeweiligen zu integrierenden Entscheidungsmodellen gefällt worden ist, so daß hier zunächst nicht mehr der Sinn der Wahl einer bestimmten Zielgröße zu erörtert zu werden braucht. Beim zweiten Wahlakt stehen grundsätzlich zwei Möglichkeiten zur Verfügung, und zwar das Streben nach einem Maximalziel oder das Ersetzen des Maximalziels durch ein leichter zu erreichendes angemessenes Ziel oder Realziel.¹

Der streng analytisch rationale Weg der Zielformulierung postuliert nur allgemeine Beurteilungskriterien wie etwa die Maximierung des Gewinns. Bei einer derartigen Vorgehensweise müssen im Entscheidungsprozeß sämtliche Interdependenzen zwischen den einzelnen Entscheidungsfeldern im Marketing berücksichtigt werden, da diese Zielformulierung dem Wesen nach simultane Entscheidungen anstreben will.² Bezüglich der Integration von Entscheidungsmodellen läßt sich folgende These aufstellen:

Eine Integration von Entscheidungsmodellen mit identischen Zielgrößen und gegebenen identischen Intensitäten, mit denen diese Zielgrößen angestrebt werden, ist nur unter Kenntnis des Interdependenzzusammenhangs zwischen den einzelnen Entscheidungsfeldern der zu integrierenden Modelle und bei Gel-

1 Vgl. Heinen, E., Das Zielsystem der Unternehmung, a.a.O., S. 82f.

2 Vgl. Wagner, H., Zielbildung und Entscheidungsprozeß bei multipolarer Struktur der betrieblichen Willensbildung, in: Sonderdruck aus: Schriften zur Kooperationsforschung, Theorie und Praxis der Kooperation A. Studien, Bd. 3, Tübingen o.J., S. 91ff, hier S. 92.

tung mindestens einer der zur Integration von Prognosemodellen genannten Bedingungen möglich.

Der zweite Weg verzichtet auf das Setzen eines Maximalzieles und verwendet unter teilweiser Vernachlässigung des Interdependenzzusammenhangs zwischen den Entscheidungsfeldern an dessen Stelle ein Realziel. Hier gilt die für Maximalziele aufgestellte These über die Integration von Entscheidungsmodellen analog.

Im folgenden soll angenommen werden, daß sich die zu integrierenden Entscheidungsmodelle durch ihre Zielfunktionen voneinander unterscheiden. Es kommt dann zu einem Zielkonflikt, wenn einerseits bestimmte Maßnahmen die Erreichung mindestens eines Ziels fördern, andererseits die Erreichung eines anderen Ziels gefährden.¹ Problem ist, ob und inwieweit sich Entscheidungsmodelle mit konfliktären Zielen integrieren lassen.

Die erste Möglichkeit zur Integration ist dann gegeben, wenn es gelingt, sämtliche Ziele der jeweils zu integrierenden Entscheidungsmodelle auf ein gemeinsames Ziel zurückzuführen. In diesem Fall hat im eigentlichen Sinn kein Zielkonflikt vorgelegen. Dem Marketing-Manager war aufgrund mangelnder Information nur nicht bewußt, daß die Verfolgung des gemeinsamen Ziels automatisch zur Verfolgung der beiden als konfliktär angesehenen Ziele führt. Hier gelten die Integrationsbedingungen für Entscheidungsmodelle mit identischen Zielgrößen analog.

Die zweite Möglichkeit besteht darin, die unterschiedliche Bedeutung einzelner Ziele möglichst aufgrund von analytischen Modellen abzuwägen, um in Anschluß daran eine Ent-

¹ Vgl. Heinen, E., Das Zielsystem der Unternehmung, a.a.O., S. 95.

scheidung zugunsten einer von drei möglichen Alternativen zu treffen.¹

Erstens: Es wird in einem Entscheidungsmodell ein Maximalziel verfolgt, während alle übrigen Ziele vernachlässigt werden. Der Zielkonflikt wird durch Unterdrückung einiger Zielgrößen gelöst. Daraus folgt, daß Modellintegrationen nicht möglich sind.

Zweitens: Es wird ein Maximalziel verfolgt, während die übrigen Ziele als Begrenzungsgrößen in das Modell eingehen.² In diesem Fall partieller Zielkonflikte separiert der Marketing-Manager zwischen Haupt- und Nebenzielen und berücksichtigt die Nebenziele als Nebenbedingungen. Es entsteht eine kombinierte Zielsetzung, für deren Einteilung in Haupt- und Nebenziele sich keine logische Begründung anführen läßt. In diesem Fall ist eine Integration von Entscheidungsmodellen unter den Voraussetzungen möglich, daß es gelingt, eine analytisch formulierbare kombinierte Zielsetzung zu finden und zumindest eine der für die Integration von Prognosemodellen notwendigen Bedingungen hier analog zutrifft.

Drittens: Werden in Entscheidungsmodellen jeweils Ziele verfolgt, die untereinander in totalem Widerspruch stehen, ist eine Modellintegration nur dann möglich, wenn es gelingt, die miteinander konkurrierenden Ziele durch ein neues Wertsystem im Sinne einer Nutzenfunktion zu ersetzen und darüber hinaus mindestens eine der zur Integration von Prognosemodellen notwendigen Bedingungen auch hier zutrifft. In diesem Fall hat sich der Marketing-Manager darum zu bemühen, an die Stelle konkurrierender Ziele ein

1 Zu dieser Vorgehensweise vgl. Meffert, H., Unternehmensziele, in: Schöttle, K.M., Hrsg., Jahrbuch des Marketing, Essen 1971, S. 22ff, hier S. 29ff.

2 Vgl. Wagner, H., Zum Problem der Zielfunktion in einer operationalen Theorie der betrieblichen Kapitaldisposition, Habilitationsschrift, Münster 1970, S. 63 (unveröffentlicht).

diesen Zielen übergeordnetes Wertesystem zu entwerfen.¹ Dieser zwar theoretisch mögliche Lösungsweg wird aber in nur wenigen Fällen zu brauchbaren Ergebnissen führen, da die den Zielen übergeordnete Nutzenfunktion für konkret-rechnerische Zwecke wenig operational sein wird.²

Bei der Integration von Prognose- und Entscheidungsmodellen liefern die Prognosemodelle den Input für die Entscheidungsmodelle. Es lassen sich ein oder mehrere Prognosemodelle mit einem Entscheidungsmodell integrieren, wenn bei den Prognosemodellen die Voraussetzungen der Prämissen-, der Variablen- und der Relationenkompatibilität erfüllt sind. Darüber hinaus muß das zu integrierende Entscheidungsmodell eine oder mehrere erklärende Variablen der Prognosemodelle enthalten. Ein Beispiel soll dieses näher verdeutlichen: In einem neoklassischen preispolitischen Entscheidungsmodell soll die Zielfunktion die Variablen Preis und Absatzmenge enthalten. Eine Integration dieses Modells mit einem Prognosemodell der Absatzmengenentwicklung ist dann möglich, wenn die Absatzmenge im Prognosemodell zu erklärende Variable ist.

Im folgenden soll der einleitend angesprochene zweite Problemkreis untersucht werden, unter welchen Gesichtspunkten eine Integration von Marketing-Modellen überhaupt sinnvoll ist.

Als wichtigstes Kriterium für den Sinn einer Modellintegration ist hier die Genauigkeit der ökonomischen Aussagen zu nennen. Werden Prognosemodelle integriert, so ist das Fehlerfortpflanzungsproblem von eminenter Bedeutung. Fehler, die in den jeweiligen Prognosemodellen vorkommen, pflanzen sich beim Ineinanderüberführen der Funktionen fort, so daß

1 Vgl. Meffert, H., Unternehmensziele, a.a.O., S. 22f.

2 Vgl. Heinen, E., Das Zielsystem der Unternehmung, a.a.O., S. 147ff.

das Modellergebnis von Integrationsstufe zu Integrationsstufe ungenauer wird.¹ Das gegenseitige Aufheben von Modell- und Datenfehlern wäre rein zufällig.² Aus diesem Gesichtspunkt ist eine Integration von Prognosemodellen nur dann sinnvoll, wenn die in jeder Integrationsstufe vorkommenden Fehler berechnet werden können.³ Dieses wird aber nur in wenigen Fällen möglich sein. Analoge Überlegungen gelten für die Integration von Entscheidungsmodellen und die Integration von Entscheidungs- und Prognosemodellen.

Ein weiterer Grund, der gegen eine Integration von Prognosemodellen spricht, ist die Möglichkeit der Ableitung rekursiver Funktionen, die ohne die rechentechnisch aufwendige Integration sehr viel genauere Ergebnisse liefern können. Schließlich ist bei zunehmender Integration damit zu rechnen, daß Multikollinearität eintritt⁴, die dazu führt, daß der Rechenaufwand sich unnötigt überproportional erhöht und die Aussagegenauigkeit des Modells beeinträchtigt wird. Darüber hinaus erweisen sich einerseits steigende Informationskosten, die auf zusätzliche Speicher- und Verarbeitungskosten zurückgeführt werden können, denen andererseits nicht exakt oder gar nicht quantifizierbare Informationserträge gegenüberstehen als Bremse bei der Integration von Prognosemodellen. Diese Feststellungen gelten für die Integration von Entscheidungsmodellen und von Prognose- und Entscheidungsmodellen analog.

-
- 1 Vgl. dazu Ludwig, R., Methoden der Fehler- und Ausgleichsrechnung, Braunschweig 1969, S.40ff.
 - 2 Koller, S. / Schaeffer, K.A., Kap.: Beurteilung der Genauigkeit, in: Stichproben in der amtlichen Statistik, hrsg. vom Statistischen Bundesamt, Wiesbaden 1960, S. 49-51, und S. 97-113.
 - 3 Vgl. dazu Ludwig, R., Methoden der Fehler- und Ausgleichsrechnung, a.a.O., S. 40f.
 - 4 Unter Multikollinearität wird die lineare Korrelation zwischen den unabhängigen Variablen einer Regressionsfunktion verstanden. Vgl. zur Multikollinearität Förster, E. / Egernmayer, F., Korrelations- und Regressionsanalyse, a.a.O., S. 210ff.

Frage ist nun, welche generellen Konsequenzen die erörterten Gesichtspunkte und Voraussetzungen der Modellintegration für den Aufbau der Modellbank eines Marketing-Informationssystem haben. Im Zusammenhang mit dieser Frage stellen sich zwei Problemkreise, und zwar erstens der mehr formelle Problemkreis, welche Reihenfolge bei der Integration von Marketing-Modellen vorzunehmen ist und zweitens der mehr materielle Problemkreis, welche Modelle für welche Marketing-Fragestellungen in die Modellbank aufzunehmen sind.

Die erste Konsequenz beim formellen Problemkreis besteht darin, die Modellbank entsprechend dem Schwierigkeitsgrad der Integration von Modelltypen aufzubauen. Tendenziell sollen zunächst Prognosemodelle aufgenommen werden, die den angeführten Integrationsbedingungen genügen. Alsdann ist als zweite Konsequenz die Integration dieser Modelle zu einem "Prognosesystem" zu versuchen. Insbesondere eignen sich hierzu die kurz- und langfristig orientierten Verfahren der Absatzprognose, wie in der Praxis entwickelte Prognosesysteme zeigen.¹ Die zu integrierenden Modelle müssen modular aufgebaut sein, d.h., daß sie sich "baukastenartig" zusammensetzen lassen. Insbesondere sind dabei die "Schnittstellen" zwischen den verschiedenen Modellen von Bedeutung, die den Output spezifizieren, der bei anderen Modellen zu Input wird. An diesen Schnittstellen ist zu prüfen, ob die genannten Integrationsvoraussetzungen gegeben sind.

Erst in der nächsten Entwicklungsphase einer Modellbank sollte die Integration von Prognose- und Entscheidungsmodellen versucht werden. Die Prognosemodelle müssen sich erst bewähren, und darüber hinaus sollten die Fehlerfort-

¹ Vgl. hierzu z.B. Lewandowski, R., Die Problematik der kurzfristigen Vorhersage und ihre Lösung mit dem Programm-System PROCOM, Manuskript des Mathematischen Beratungs- und Programmierdienst G.m.b.H., Dortmund o.J., S. 32ff.

pflanzungsmöglichkeiten in diesen Modellen bekannt sein, damit diese von Prognose- in Entscheidungsmodelle verhindert werden. Erst im fortgeschrittenen Entwicklungsstadium einer Modellbank sollte ein Integrationsversuch von Entscheidungsmodellen vorgenommen werden, da sich diese Modelle aufgrund der abgeleiteten Gesichtspunkte am schlechtesten integrieren lassen.

Der formelle Problemkreis steht in interdependentem Zusammenhang mit dem materiellen Problemkreis, welche Modelle für welche Marketingfunktionen in die Modellbank aufgenommen werden sollen.¹ Beide Problemkreise lassen sich nicht exakt ohne Bezug auf tatsächlich angewendete Modelle und konkrete Gegebenheiten eines vorliegenden Falles beantworten. Es soll hier nur festgestellt werden, daß man am besten so vorgehen wird, daß zunächst ein globales Marketing-Ziel formuliert wird. Liegt dieses Ziel fest, so besteht die weitere Aufgabe des Marketing-Management darin, aus diesem Ziel adäquate Unterziele für die einzelnen Funktionsbereiche des Marketing abzuleiten. Alsdann wird gefragt, auf welchen zentralen Variablen die einzelnen Funktionsziele basieren. Nach Identifizierung dieser zentralen Variablen wird dann versucht, ihre Entwicklung mittels Prognosemodellen vorherzusagen. Die sich anschließende weitere Entwicklung stimmt dann mit der skizzierten Vorgehensweise entsprechend dem Schwierigkeitsgrad der Integration von Modelltypen überein.

¹ Vgl. dazu den Punkt 1.2 dieses Abschnitts.

4. Schlußfolgerungen für die Modellselektion und -entwicklung

Aus dem theoretisch-intersubjektiven und dem benutzerorientiert-subjektiven Ansatz ergeben sich im Zusammenhang mit den erörterten Kriterien für die Modellselektion und -entwicklung Konsequenzen für Marketing-Modelle in einer Modellbank. Es soll hier eine abschließende Würdigung beider Ansätze vorgenommen und versucht werden, darauf aufbauende Thesen über die Aufnahme von Modellen in eine Modellbank zu formulieren.

Der theoretisch-intersubjektive Ansatz geht davon aus, daß das Problem der Selektion und Entwicklung von Marketing-Modellen nur unter dem Primat der homomorphen Abbildung von Marketing-Phänomenen in Modellen gelöst werden kann. Gelingt es einem Modellkonstrukteur, eine homomorphe Abbildung eines Realsystems zu erhalten, dann ist dieses Modell für die Aufnahme in die Modellbank geeignet. Die tatsächliche Übereinstimmung zwischen Realität und Modell kann bei diesem Ansatz mittels Abbildungskalkülen geprüft werden.¹ Darüber hinaus zeichnet sich der theoretisch-intersubjektive Ansatz durch starken Reduktionismus aus. Dieses besagt, daß eine Fülle von an Modelle zu stellende Anforderungen, die nicht benutzerorientiert sind, vom Kriterium der Homomorphie umfaßt werden. Ist Homomorphie gegeben, sind ebenfalls Kriterien wie z.B. Genauigkeit und Validität gegeben. Dieser Reduktionismus ist gleichermaßen von Vorteil wie auch von Nachteil. Der Vorteil besteht darin, daß bei gegebener Homomorphie nicht mehr geprüft zu werden braucht, ob objektive Anforderungen an das Modell gegeben sind; sie sind immer erfüllt. Dieser Vorteil kann aber nur über eine Reihe von Nachteilen "erkauft" werden, die sich wiederum aus dem starken Reduktionismus dieses Ansatzes ergeben.

¹ Vgl. dazu z.B. Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 266ff.

Soll ein System homomorph in einem Modell abgebildet werden, so muß der theoretisch-intersubjektive Ansatz dieses System streng axiomatisch abgrenzen und dann abbilden, um den Homomorphiegrad des Modells konkret-rechnerisch überprüfen zu können.¹ Mit dieser notwendigen Axiomatisierung ist gleichzeitig eine zunehmendere Realitätsferne verbunden, die sich u.a. aufgrund der bei der Systemabgrenzung notwendigen Prämissen sachlicher als auch abbildungsmathematischer Art ergibt. Daraus folgt, daß der theoretisch-intersubjektive Ansatz nur in den Grenzen des Realsystems die Homomorphie analysieren kann, von denen bereits bei der Modellbildung angenommen wurde, daß sie für den Homomorphiegrad des Modells entscheidend sind. Dieses streng rationale Vorgehen impliziert weiterhin, daß Konsens über die Systemziele besteht und die Mittel zur Zielerreichung a-priori bekannt sind, so daß das Modell den einzigen Zweck verfolgt, aus einem Satz von endlichen Alternativen mit ihren sämtlichen Konsequenzen eine Lösung zu ermitteln. Verhaltenswissenschaftliche Phänomene, das Problem der Unsicherheit der Datenentwicklung, benutzerorientierte Gesichtspunkte, kurzum sämtliche Phänomene, die sich abbildungsmathematisch nicht streng axiomatisieren lassen, müssen vom theoretisch-intersubjektiven Ansatz in seinen Homomorphiekalkülen vernachlässigt werden.

Die Vorgehensweise dieses Ansatzes ist vereinfachend ausgedrückt auf die Anwendung derjenigen Methoden ausgerichtet, bei denen sich bei Einsatz geeigneter Daten und bei Angabe der Ziele eine Lösung quasi von selbst ergibt. Aus diesen Gründen ist der theoretisch-intersubjektive Ansatz nur bedingt geeignet. Hilfen zur Selektion und Entwicklung von Marketing-Modellen zu liefern. In der Anwendung dieses Ansatzes ist ein Grund für das Versagen der bisherigen Bemühungen zur Entwicklung benutzerorientierter Marketing-Mo-

¹ Vgl. Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, a.a.O., S. 269-274.

delle zu sehen.¹ Er ist zwar theoretisch konsistent, weist aber aufgrund der Komplexität seiner Formulierbarkeit Verständnisprobleme für das Marketing-Management auf. Somit kann die erste These für die Modellselektion und -entwicklung nur lauten:

Das Homomorphieprinzip darf nicht zum generellen Entscheidungskriterium der Modellselektion und -entwicklung gemacht werden. Es ist insbesondere durch benutzerorientierte Anforderungen an die Modelle zu ergänzen.

Darüber hinaus müssen für bestimmte Marketing-Funktionen konzipierte Modelle ggf. unter Berücksichtigung spezifischer unternehmensexterner und -interner Besonderheiten modifiziert werden. Dieses kann einerseits in der fehlenden oder mangelhaften Beschaffenheit von Daten, andererseits in den unrealistischen Modellprämissen begründet sein.

Dem Gesichtspunkt, daß mittels Homomorphie-Kalkülen zwar Modelle entwickelt und selektiert werden können, die aufgrund ihrer Axiomatik völlig logisch sind, es aber nicht gestatten, empirische Marketing-Phänomene anwendungsorientiert zu berücksichtigen, wird durch den Decision-Calculus-Ansatz Rechnung getragen. Darin liegt der große Verdienst dieses Ansatzes. Der auf dem Gedanken des Problemlösungsprozesses basierende Decision Calculus weist durch den möglichen Zusammenhang zwischen den an der Modellselektion und -entwicklung Beteiligten und ihren kognitiven Prozessen Vorteile gegenüber der Vorgehensweise des theoretisch-intersubjektiven Ansatzes auf. Problemstrukturierung und Lösungssuche bedingen umfangreiche Lernprozesse bei den daran Beteiligten, die durch phasenspezifische Informationen² in-

¹ Vgl. zu diesbezüglichen wissenschaftstheoretischen Gesichtspunkten Bochenski, I.M., Die zeitgenössischen Denkmethoden, a.a.O., S. 49.

² Vgl. dazu den Punkt 2.2 dieses Abschnitts.

duziert werden. Dieser auf Problemstrukturierung und Lösungssuche ausgerichtete problemspezifische Lernprozeß ist bei den an der Selektion und Entwicklung von Marketing-Modellen Beteiligten auf den Aufbau genereller Verfahrensmuster ausgerichtet. Derartige "Erfahrungskomplexe" lösen bei den Beteiligten eine bestimmte "wenn-dann"-Reaktion bei der Problemlösung aus. Wenn z.B. ein Beteiligter ein bestimmtes Phänomen bei der Problemstrukturierung beobachtet, dann weiß er aus Erfahrung, welche Auswirkungen dieses für seine Rolle, die Rollen anderer Beteiligter und für das zu selektierende oder zu entwickelnde Modell hat.

Unabhängig von der Struktur der Marketing-Organisation lassen sich bei der Selektion und Entwicklung von Marketing-Modellen eine Reihe von spezifischen informationsverarbeitenden Rollen unterscheiden, und zwar die des Marketing-Managers, die des Systemanalytikers, die des Computerfachmanns und die des Marktforschers. Diese Rollen sollen hinsichtlich ihrer Aufgabendifferenzierung bei der Lösung eines Problems analysiert werden, um weitere Thesen über die Aufnahme von Modellen in die Modellbank eines Marketing-Informationssystems abzuleiten.

Bei der Lösung eines Marketing-Problems müssen bestimmte Aspekte operationalisiert werden, so daß die Übereinstimmung zwischen dem Problemlösungskonzept und der Wirklichkeit möglichst groß ist. Diese von einem internen Umweltmodell ausgehende begriffliche Zerlegung eines Aspektes ist eine Aufgabe, die allein dem Systemanalytiker überlassen bleibt. Um den Zusammenhang mit dem zu lösenden Problem nicht zu verlieren, wurden mathematisch formulierte Umweltmodelle entwickelt, die statt gemeinsamer interner Umweltmodelle zwar dem methodischen Vorgehen des Systemanalytikers und des Computerfachmanns entsprechen, die aber i.d.R. vom Marketing-Manager in Ermangelung einer gemeinsamen Sprache zur Kommunikation nicht voll verstanden werden. Der Marketing-Manager befaßt sich primär mit Pro-

blemen, deren Komplexität und scheinbare Ungeordnetheit sich dem Systemanalytiker und Computerfachmann nicht vermitteln lassen. So ist es zu erklären, daß sich Spezialisten oft um der Lösbarkeit willen irrelevanten Problemen zuwenden, nur weil sie lösbar sind.¹ Der Marktforscher hat primär die Aufgabe, Daten zu sammeln und aufzubereiten. Seine Rolle steht in gewissem Widerspruch zu der des Systemanalytikers und des Computerfachmanns, da für ihn Strukturierungsprobleme nicht dominant sind.

Selbst wenn der Marketing-Manager bei der Selektion und Entwicklung von Modellen hinzugezogen wird, deren Daten vom Marktforscher beschafft wurden, so wird er trotz des Verständnisses der vom Systemanalytiker und Computerfachmann vorstrukturierten Probleme nur selten in die Gründe vordringen, wo die entscheidenden Prämissen für die Abbildung der Wirklichkeit im Modell liegen.²

Diese Feststellung gestattet es, folgende These über die Entwicklungsrichtung von Modellen aufzustellen:

Modelle und Verfahren aus der Methodenbank müssen sinnvoll aufeinander abgestimmt sein, so daß sämtliche Rollenträger die Aussagefähigkeit, die Prämissen und die Anwendungsgrenzen von Modellen und Verfahren kennen.³

Diese These besagt, daß beispielsweise erst Verfahren vom Marktforscher eingesetzt werden, mittels deren Daten aufbereitet werden können, die dann als Input in komplexeren Modellen Anwendung finden. Sie kann als Dominanzprinzip statistischer multivariater Datenanalysemethoden über

1 Vgl. Luhmann, N., Zweckbegriff und Systemrationalität, a.a.O., S. 221.

2 Vgl. Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, a.a.O., S. B-467.

3 Vgl. Montgomery, D.B. / Urban, G.L., Marketing Decision-Information Systems, a.a.O., S. 230.

Marketing-Modelle im Aufbauprozeß eines Marketing-Informationssystems verstanden werden. In dieser These kommt der Grundsatz des schrittweisen Vorgehens beim Aufbau von Marketing-Informationssystemen und die Notwendigkeit eines damit korrelierten Verständnisses aller Rollenträger zum Ausdruck.

Der Computerfachmann verfolgt bei der Selektion und Entwicklung von Marketing-Modellen Ziele, die weitab vom Interessensgebiet des Marketing-Managers liegen. Er sieht sein Ziel darin, möglichst umfassende Softwarepakete zu entwickeln. Darüber hinaus ist er bestrebt, möglichst effiziente Speicher- und Zugriffstechniken anzuwenden, ohne oft danach zu fragen, ob der Marketing-Manager diese zu seiner Arbeit benötigt. Der Marketing-Manager erhofft sich dagegen ein Instrumentarium, das es ihm ermöglicht, die Komplexität zu strukturieren. Diese Gesichtspunkte ermöglichen es, folgende These aufzustellen:

Wenn komplex-strukturierte Marketing-Modelle eingesetzt werden, entstehen Schwierigkeiten bezüglich der Zugriffsmöglichkeiten und der Zugriffszeiten bei den in diese Modelle eingehenden Daten. Darüber hinaus können Verständnisprobleme seitens des Marketing-Managers eintreten.

Diese These soll zum Ausdruck bringen, daß eine Modellbank, die Modelle unterschiedlichen Abstraktionsgrades enthält, auch Angaben über Zugriffs- und Modellaufzeiten enthalten muß. Letztlich ist es eine Kosten- und daraus abgeleitet eine Speicherkapazitätsfrage, ob und in welchem Umfang Programme on-line, d.h. für den Marketing-Manager direkt abrufbar, gespeichert werden können. Von der Art der Speicherung und den gespeicherten Programmen hängt die Qualität und die Brauchbarkeit eines Marketing-Informationssystems im wesentlichen ab.¹

¹ Vgl. Nieschlag, R. / Dichtl, E. / Hörschgen, H., Marketing, 4. Aufl., Berlin 1971, S. 591f.

Der Marktforscher ist schließlich mit der Beschaffung von Daten betraut, ohne daß er oft im Detail die beabsichtigte Verwendung durch den Marketing-Manager und die Anforderungen seitens des Computerfachmanns an diese Daten kennt. Ihn tangiert primär das Problem der Wirtschaftlichkeit der Datenbeschaffung. In diesem Zusammenhang läßt sich folgende These aufstellen:

Es ist sinnvoll, in die Modellbank nur solche Modelle aufzunehmen, für die in der Datenbank bereits Daten gespeichert sind oder durch Anpassung der Datenbank ohne überproportionale Kosten im Verhältnis zu den erwarteten Informationserträgen gewonnen werden können.¹

Diese These ist unter Wirtschaftlichkeitsgesichtspunkten einleuchtend. Es ist beispielsweise verfehlt, komplexe Simulationsmodelle in die Modellbank aufzunehmen, wenn die Datenbank nicht die dazu notwendigen Daten liefern kann. Darüber hinaus ist bei der Modellaufnahme in die Modellbank zu beachten, daß sich in der für die Implementierung des Modells notwendigen Daten solche unterschiedlichen Genauigkeitsschichten befinden. Daraus folgt, daß die Modelle in Hinblick auf diese Genauigkeitsschichten ausgerichtet sein sollten. Da durch die in den Modellen stattfindenden Transformationsprozesse Fehlerfortpflanzungen stattfinden können, oder es aber zu Fehlerkumulierungen kommen kann, ist den Prinzipien der Ungewissheitsreduktion und -absorption besondere Aufmerksamkeit zu schenken.²

1 Die hier fast "trivial" erscheinende Feststellung gewinnt durch die Problematik des Kosten-Nutzen-Vergleichs erheblich an Bedeutung. Zur Informationskostenrechnung vgl. Wild, J., Informationskostenrechnung auf der Grundlage informationeller Input-, Output- und Prozeßanalysen, in: Zeitschrift für betriebswirtschaftliche Forschung, 22. Jg., 1970, S. 218ff und Hirsch, R.E., Informationswert und -kosten und deren Beeinflussung, a.a.O., S. 670ff.

2 Vgl. Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, a.a.O., S. 32 und S. 52.

Die Vorgehensweise des Decision Calculus besteht u.a. darin, daß im Zusammenhang mit Problemlösungsprozessen auf ein heuristisches, zu neuen Lösungen führendes Vorgehen abgestellt wird. Durch eine überwiegend zeitlich festgelegte Reihenfolge von einzelnen Phasen innerhalb eines Problemlösungsprozesses können Marketing-Manager und andere an der Modellselektion und -entwicklung Beteiligte sich im Rahmen ihrer Rolle auf diese Phasen spezialisieren. Darüber hinaus können Marketing-Manager, Systemanalytiker, Computerfachleute und Marktforscher innerhalb einer Klasse von Problemen durch Interaktion bei der Problemstrukturierung und Lösungssuche ihre Erfahrungen im Sinn eines gemeinsamen informellen Lösungsprogramms aufbauen. In Kontext mit einem derartigen Programm kann die Informationsverarbeitungs-Kapazität erweitert und mehr Information im Lauf einer Modellselektion bzw. -entwicklung verarbeitet werden.

Die Modellstrukturen haben durch die Fortschritte bei der Entwicklung mathematischer Modelle für Marketing-Funktionsbereiche eine erhebliche Änderung erfahren. Obgleich z.B. marginalanalytische und lineare Optimierungsverfahren hinsichtlich ihrer Anwendbarkeit modifiziert wurden, entsprechen die dort abgeleiteten Lösungen nicht der Denkweise eines Marketing-Managers. Er wird sie deshalb mit der Begründung unrealistischer Prämissen ablehnen. Darüber hinaus wird er nur selten in der Lage sein, unrealistisch erscheinende Annahmen der Modelle so zu modifizieren, daß sein internes Umweltmodell von dem zu lösenden Problem z.B. mit marginalanalytischen und linearen Optimierungsverfahren kompatibel ist.

"Diese Ausführungen machen .. deutlich, warum die Entwicklung der modernen Optimierungstechniken wie die lineare, nichtlineare oder dynamische Programmierung nicht die vielfach erwartete wesentliche Verbesserung in der Entwicklung von Marketing-Modellen brachte. Denn das Problem der Lö-

sung von Führungsproblemen des Marketing besteht weniger darin, effiziente mathematische Algorithmen zu finden, sondern es besteht primär in der Strukturierung und Ordnung des Entscheidungsproblems und in der simultanen Verarbeitung aller notwendigen Daten zu Entscheidungshypothesen, deren Akzeptierung letztlich von der Risikopräferenz eines Managers subjektiv bestimmt wird."¹ Aus diesen Gründen besteht gegenwärtig die Tendenz zum Übergang auf heuristische Verfahren der Problemlösungstheorie. Mittels dieser neueren Ansätze kann insbesondere dem im Marketing vorhandenen Interdependenzproblem besser Rechnung getragen werden als mit den herkömmlichen analytischen Optimierungstechniken.

Unter diesem Aspekt der Problemlösung ist der Decision Calculus zu sehen und erlangt aus dieser Warte seinen Stellenwert. In seiner problemlösungsorientierten Interpretation liefert er zwar neue Denkanstöße, die aber nicht zwingend dazu führen müssen, den theoretisch-inter-subjektiven Ansatz völlig zu verdrängen. Keinem dieser beiden Ansätze kommt ein absoluter Prioritätsanspruch zu, da beide Ansätze gewisse Vor- und Nachteile besitzen. So dominiert beim theoretisch-inter-subjektiven Ansatz die Kalkülisierbarkeit von abgebildeten Phänomenen, während der benutzerorientiert-subjektive Ansatz (Decision Calculus) mehr in Richtung auf ein heuristisch orientiertes Vorgehen zielt.

¹ Krautter, J., Marketing-Entscheidungsmodelle, a.a.O., S. 31f. (Im Original teilweise hervorgehoben.)

V. Ausblick auf künftig mögliche Forschungsstrategien
bei Marketing-Informationssystemen

Der Ansatz von LITTLE hat gezeigt, daß nur durch ein Abstellen auf kognitive Problemlösungsprozesse der an der Entwicklung und Benutzung eines Marketing-Informationssystems Beteiligten geeignete managementorientierte Modelle selektiert und entwickelt werden können. Dieser Gedanke wird ebenfalls von SZYPERSKI¹ vertreten, der feststellt, daß sich empirisch-kognitives und real-konstruktives Wissen im Wissenschaftler vereinigen müssen, um Forschung auf dem Gebiet der Marketing-Informationssysteme betreiben zu können. Obwohl LITTLE seinen Ansatz auf den Problemkreis der Modellelektion und -entwicklung beschränkt hat, ist dieser Ansatz doch derart generell, daß es geeignet erscheint, ihn teilweise zur Grundlage eines Ausblicks auf künftig mögliche Forschungsstrategien zu machen. Darüber hinaus wird LITTLESches Gedankengut von MASON und MITROFF² aufgegriffen, synthetisiert und zu einem programmatischen Imperativ erweitert.

MASON und MITROFF gehen bei ihrem Forschungsprogramm von der Definition eines Informationssystems aus. Sie postulieren, daß ein Informationssystem zumindest aus einer Person eines gewissen psychologischen Typs besteht, die ein Problem innerhalb des organisatorischen Kontexts lösen will, für das sie eine Erkenntnismethode benötigt, wo bei ihr die Evidenz der Lösung durch irgendeine Darstellungsart bewußt gemacht wird.³ Das Neue dieses Ansatzes besteht darin, daß die fünf Komponenten obiger Definition als Variablen angesehen werden, die es zu analysieren gilt, um die Bereiche gegenwärtiger und künftiger Forschungsrichtungen abzustecken.

1 Vgl. Szyperski, N., Gegenwärtiger Stand und Tendenzen der Entwicklung betrieblicher Informationssysteme, in: IBM Nachrichten, 23. Jg., 1973, S. 473ff, hier S. 482.

2 Die folgenden Gedankengänge lehnen sich an die Arbeit von Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, in: Management Science, Vol. 19, 1973, S. B-475ff an.

3 Vgl. ebenda, S. 475.

Es wurde bei der Erörterung des Decision-Calculus festgestellt, daß sich bei der Selektion und Entwicklung von Marketing-Modellen vier Rollenträger unterscheiden lassen, die spezifische informationsverarbeitende Rollen zu erfüllen haben. Diese spezifischen Rollen gelten für den Aufbau eines Marketing-Informationssystems analog. Der Marketing-Manager, der Systemanalytiker, der Computerfachmann und der Marktforscher haben bei der Entwicklung eines Marketing-Informationssystems die Aufgabe, ihr empirisch-kognitives und ihr real-konstruktives Wissen zweckorientiert zu integrieren. Der Bereich des empirisch-kognitiven Wissens wird von MASON und MITROFF¹ durch die These abgedeckt, daß insbesondere der Marketing-Manager eine Erkenntnismethode benötigt, die auf seine rollenspezifischen Probleme zugeschnitten ist und nicht auf die der übrigen an der Entwicklung eines Marketing-Informationssystems Beteiligten. Da der Marketing-Manager solche Informationen benötigt, die mit seinem internen Umweltmodell kompatibel sind und seiner Psychologie entsprechen, ist es nicht nur Aufgabe des Systemanalytikers, des Computerfachmanns und des Marktforschers herauszufinden, welche Informationen der Manager tatsächlich benötigt, sondern auch welche Darstellungsform der Information unter Berücksichtigung des psychologischen Typs des Managers geeignet ist.² Der Bereich des real-konstruktiven Wissens umfaßt dagegen u.a. die Entwicklung von Gestaltungsprinzipien für die Daten-, Methoden- und Modellbank.

Aufgrund der Rollenerwartungen und der damit zu unterstellenden psychologischen Kategorien der an der Entwicklung eines Marketing-Informationssystems Beteiligten, nehmen MASON und MITROFF³ eine Klassifizierung vor. Sie basiert

1 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 457.

2 Vgl. ebenda.

3 Vgl. ebenda, S. 476-479.

auf den zwei dominanten psychologischen Funktionen, nämlich Wahrnehmung und Bewertung, deren sich ein Individuum bedient, um reale Objekte zu begreifen. Die Wahrnehmung kann sensitiv oder intuitiv erfolgen, während eine Bewertung durch Denken oder Empfinden vorgenommen werden kann. Dieses gestattet es, folgende psychologische Klassifizierung zu entwerfen:

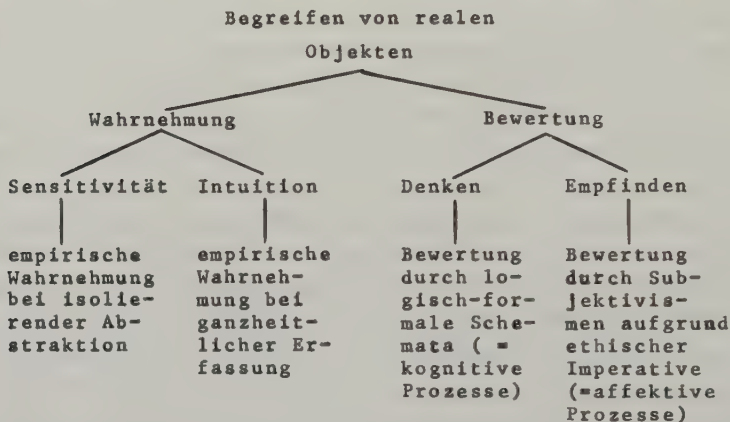


Abb. 20: Begreifen von realen Objekten

Das Begreifen von realen Objekten oder Phänomenen der Realität innerhalb des Wahrnehmungsprozesses geschieht beim sensitiven psychologischen Typ aufgrund isolierender Abstraktion der Realität. Er versucht diese Phänomene zu analysieren, zerlegt sie dann in ihre einzelnen Elemente und synthetisiert sie schließlich, um zu einer Feststellung zu gelangen. Ein typisches Beispiel einer derartigen Vorgehensweise ist der in dieser Arbeit dargestellte sprachanalytische Ansatz. Im Gegensatz dazu versucht die intuitive Vorgehensweise, Objekte der Realität ganzheitlich zu erfassen und begreift sie als Möglichkeiten.¹

¹ Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 477.

Der Bewertungsprozeß von realen Objekten, von dem angenommen wird, daß er unabhängig vom Wahrnehmungsprozeß ist, ist durch Denken und Empfinden möglich. Die Bewertung durch Denken erfolgt aufgrund in sich geschlossener logisch-formaler Schemata (Denkgesetze), die darüber Aufschluß geben, ob die vorgenommene Bewertung richtig oder falsch ist. Im Gegensatz zur abstrakten Bewertung durch Denken zeichnet sich die Bewertung durch Empfinden durch konkrete Subjektivismen des einzelnen Bewertenden aus, wobei diese Subjektivismen aufgrund ethischer Imperative gebildet werden.

Von ihren Rollenerwartungen her lassen sich nun die an der Entwicklung und Benutzung eines Marketing-Informationssystems Beteiligten in die genannten psychologischen Kategorien einteilen, unbeschadet der Tatsache, ob sie effektiv einer bestimmten Kategorie angehören oder nicht. MASON und MITROFF¹ stellen dazu fest, daß trotz einer Vielfalt möglicher psychologischer Konfigurationen eine Einteilung der an der Entwicklung eines Informationssystems Beteiligten als "Karikaturen" den Vorteil hat, daß das entwickelte Schema heuristischen Wert besitzt, um Persönlichkeiten und Management-Stile zu charakterisieren.

Beim Wahrnehmungsprozeß sind nach MORRIS² strategische Planungsprozesse mehr der intuitiven Vorgehensweise vorbehalten, die darauf abzielt, Marktchancen und Zukunftsentwicklungen intuitiv zu erfassen. Bei den operativen Prozessen dominiert dagegen das sensitive Vorgehen, weil es hierbei besonders auf quantitative Situationsanalysen ankommt.

1 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 478.

2 Vgl. Morris, W.T., Intuition and Relevance, in: Management Science, Vol. 14, 1967, S. B-157ff, hier S. 158ff.

Aufgrund seiner Rollenerwartung wird sich der Marketing-Manager mehr in die Kategorie des intuitiv Vorgehenden einordnen lassen, weil bei Führungsentscheidungen die Kreativkomponente eine nicht zu vernachlässigende Rolle spielt. Der Systemanalytiker und der Computerfachmann können aufgrund ihrer jeweiligen Rollenerwartung in die Kategorie der sensitiv Vorgehenden eingeordnet werden, da von ihnen verlangt wird, daß sie streng rational die ihnen vorliegenden Probleme analysieren. Schließlich kann man dem Marktforscher eine dualistische Stellung zwischen sensitivem und intuitivem Vorgehen einräumen. Die Sensitivität wird bei ihm dann dominieren, wenn es darum geht, Daten zu analysieren und mittels statistischer Methoden aufzubereiten und auszuwerten. Intuitives Vorgehen wird dann für ihn von Bedeutung, wenn es darum geht, ein komplexes Marktforschungsproblem in seiner Ganzheit zu erfassen, um Lösungsvorschläge zu unterbreiten.

Bei der Bewertung der wahrgenommenen Objekte der Realität ist die Bewertung durch Denken beim Marketing-Manager einerseits dann gegeben, wenn z.B. ein Entscheidungsproblem mit einem formal-rationalen Modell der Bewertung von Alternativen gelöst werden soll.¹ Andererseits wird er vielfach aufgrund mangelnder Wohlstrukturiertheit seiner Entscheidungsprobleme dazu gezwungen sein, Entscheidungen auf Judgement-Basis zu fällen, wobei ihm seine akkumulierten Erfahrungen wertvolle Hilfe leisten.

Der Zusammenhang zwischen dem Decision-Calculus von LITTLE und dem auf Forschungsstrategien bei Informationssystemen bezogenen Ansatz von MASON und MITROFF wird besonders dadurch deutlich, daß jeder psychologische Typ in der genannten Systematik einen unterschiedlichen Informationsbegriff hat, der von Bedeutung beim Aufbau eines Marketing-

¹ Vgl. Mason, R.O. / Mitroff, I.I., A Programm for Research on Management Information Systems, a.a.O., S. 478.

Informationssystem¹ ist. Die im Zusammenhang mit der phasenspezifischen Verwendung der Information bei den verschiedenen Stufen des Problemlösungsprozesses gewonnenen Erkenntnisse² lassen sich auf die Rollen der an der Entwicklung eines Informationssystems Beteiligten übertragen. Für den Manager ist Information zweckorientiertes Wissen, das ihn befähigt, seine Entscheidungen zu treffen. Für den Systemanalytiker und den Computerfachmann ist Information das, was systematisiert und analysiert werden kann und sich aus der Verarbeitung von Daten ergibt. Schließlich ist Information für den Marktforscher das Äquivalent seiner beschafften oder zu beschaffenden Daten.

An die Untersuchung möglicher psychologischer Typen schließt sich die Analyse der zweiten Komponente der Informationssystem-Definition von MASON und MITROFF - die Strukturierung des zu lösenden Problems - an. Die Autoren unterscheiden zwischen wohlstrukturierten und schlechtstrukturierten Entscheidungsproblemen. Nach Maßgabe der Unsicherheit über zu wählende Alternativen übernehmen sie die Einteilung in Entscheidungen unter Sicherheit, Entscheidungen unter Risiko und Entscheidungen unter Unsicherheit. Im Gegensatz zu strukturierten Entscheidungsproblemen zeichnen sich schlechtstrukturierte Entscheidungen primär dadurch aus, daß der Marketing-Manager nur aufgrund seines Erfahrungskomplexes gute Entscheidungen treffen kann. "Der gegenwärtige Design von MIS hat sich fast ausschließlich mit der Beschaffung von Information für wohldefinierte und wohlstrukturierte Probleme befaßt. Relativ wenig Arbeit wurde bei Informationssystemen vorgenommen, die sich mit der Verbesserung des Schätzproblems und der Behandlung unstrukturierter und verzwickter Probleme befassen."³

1 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 478.

2 Vgl. dazu den Punkt 2.2 dieses Abschnitts.

3 Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 480 (Übersetzung vom Verfasser).

Es liegt der Schluß nahe, daß Information aus einem Marketing-Informationssystem - wie immer es strukturiert ist und wie umfangreich auch immer der Datenbestand ist, der in seiner Datenbank vorhanden ist - nur eine ergänzende Funktion für Entscheidungsprozesse im Marketing haben kann. Damit soll nicht bestritten werden, daß Daten und methodisch orientierte Informationsverarbeitung einen Stellenwert bei Entscheidungsprozessen besitzen und im Sinne einer Unterstützungsfunktion des Marketing-Management verbessert und ausgebaut werden müssen. Diesbezügliche Vorgehensweisen legen es nahe, die Informationsverarbeitung in allen Phasen mit nachweislich brauchbaren Daten und in einer anwendbaren, je nach dem Fortschritt der Lösungssuche aktuellen Verarbeitungsform zu unterstützen. Welche Daten für welche Marketing-Funktionen brauchbar sind, und welche Hilfe mittels Computerprogrammen daraus gewonnen werden können, muß im voraus bestimmt werden. Durch diesen Zwang zur Vorherbestimmung sind derartige Konzepte nur bei Marketing-Funktionen realisierbar, die sich durch ein routinemäßiges Muster der Informationsverarbeitung kennzeichnen lassen. Eine einzelne Aufgabe wird hier nicht als einmalig angesehen, sondern soweit wie möglich als Teil von Aufgaben einer gegebenen Klasse. So kann von Aufgabe zu Aufgabe Zeit gewonnen werden, um die den Daten zugrundeliegenden operationalen Definitionen auf ihre inhaltliche Übereinstimmung mit den unterschiedlichen Marketing-Funktionen zu überprüfen, um dann mit den tatsächlich geeigneten Daten Auswertungen vorzunehmen. Diese Auswertungen lassen dann i.d.R. eine Anzahl von ausreichend datenmäßig abgesicherten Schlüssen zu. Die Erfahrung von Marketing-Managern ist bei diesem Prozeß bewußt zu berücksichtigen.

Die nächste Komponente der Definition eines Informationssystems nach MASON und MITROFF¹ ist die des organisatorischen Kontext, in dem Probleme gelöst werden. Setzt man

¹ Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.o., S. 483f.

Informationssystem und Organisationsstruktur gleich, so stehen beim Informationssystem die Informationsquellen und -arten im Vordergrund, die mit dieser Organisationsstruktur zusammenhängen. Die Organisationsstruktur betrachtet die Ebenen der Probleme, die die Organisation zu lösen hat. MASON und MITROFF¹ unterscheiden hier strategische Planung, Probleme der Managementebene und Probleme der operativen Ebene.

Die gegenwärtig entwickelten Informationssysteme befassen sich überwiegend mit Problemen auf operativer Ebene. Diese Arten von Informationssystemen stehen aber in Widerspruch zu Organisationsstrukturen, die für die Zukunft am besten geeignet sind. Insbesondere wurden bei der gegenwärtigen Entwicklung zwei Aspekte vernachlässigt, und zwar erstens der Einfluß der Organisationsstruktur auf den Design des Informationssystems und zweitens der Einfluß des Informationssystems auf die organisatorischen Rollen.² Diese Aspekte des organisatorischen Kontext sind noch nicht systematisch erforscht worden, und zwar besonders nicht im Zusammenhang mit anderen Aspekten des Aufbaus eines Informationssystems.³

Die nächste zu untersuchende Komponente der Informationssystem-Definition ist die der Erkenntnismethode oder des wissenschaftstheoretischen Vorgehens. Nach der agnostischen Auffassung über wissenschaftstheoretische Vorgehensweisen unterscheiden MASON und MITROFF⁴ in Anlehnung an CHURCHMAN⁵ fünf mögliche Ansätze, und zwar den empirisch-

1 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 483.

2 Vgl. ebenda, S. 484.

3 Vgl. ebenda.

4 Vgl. ebenda, S. 480ff.

5 Vgl. Churchman, C.W., The Design of Inquiring Systems, New York 1971, S. 14ff.

induktiven, den rational-deduktiven, den kritizistisch-subjektiv-idealistischen, den dialektisch-objektiv-idealistischen Ansatz und den kybernetisch-lernorientierten Ansatz.

Der empirisch-induktive Ansatz ist durch den Schluß vom Speziellen zum Allgemeinen gekennzeichnet. Er geht von einer Menge elementarer empirischer Beobachtungen und Urteilen aus und konzentriert sich auf die Entwicklung eines Systems immer allgemeiner werdenden Aussagen. Dieser Ansatz ist nach Auffassung der Autoren für gutstrukturierte Probleme geeignet, bei denen Konsens über die Art der Probleme bei den Beteiligten besteht. Der Aufbau von Datenbanken und die Entwicklung der meisten statistischen Verfahren basiert auf diesem wissenschaftstheoretischen Ansatz.

Beim rational-deduktiven Ansatz wird vom Allgemeinen auf das Spezielle geschlossen. Informationen, die aus Modellen gewonnen werden, die einer axiomatisch-symbolischen Darstellung genügen, entsprechen diesem Ansatz. Sein Vorteil besteht darin, daß genau spezifiziert werden kann, was als Beweis der abgeleiteten Vorschläge oder Theoreme dienen soll. Weitere Vorteile bestehen darin, daß diesbezügliche Modelle sich durch Konsistenz, Vollständigkeit und Verständlichkeit auszeichnen. Der rational-deduktive Ansatz eignet sich für gut-strukturierte Probleme, für die eine analytische Formulierbarkeit mit Lösungsgarantie existiert.

Der kritizistisch-subjektiv-idealistische Ansatz geht davon aus, daß ein a-posteriori ein a-priori voraussetzt. So setzt z.B. die Datensammlung (a-posteriori) eine Theorie (a-priori) voraus, unbeschadet der Tatsache, wie implizit und informal diese Theorie ist. Die Grundfragestellung dieses Ansatzes besteht darin, wie synthetische Urteile, d.h. Urteile aufgrund empirischer Wahrnehmung, a-priori möglich sind. Um diese Frage zu beantworten, müs-

sen mindestens über ein zu erklärendes Phänomen zwei alternative Theorien existieren. Die Hoffnung dabei ist, daß eine dieser zwei alternativen Theorien geeignet ist, das Problem eines Entscheidungsträgers zu erklären. Der kritizistisch-subjektiv-idealistische Ansatz ist für Systeme typisch, die mittels unterschiedlicher Modelle abgebildet werden können, wobei dann zu entscheiden ist, welche Theorie den höchsten Erklärungswert besitzt. Während der empirisch-induktive und der rational-deduktive Ansatz ein Phänomen nur aus einer Perspektive sehen, ermöglicht der kritizistisch-subjektiv-idealistische Ansatz aufgrund der notwendigen Theorienpluralität eine Erklärung vieler Perspektiven des Problems. Da dieser Ansatz weder empirisch noch rational ist, eignet er sich zur Behandlung "gemäßigt-schlechtstrukturierter" Entscheidungen.

Die Vorgehensweise des dialektisch-objektiv-idealistischen Ansatzes besteht in einer vollständig antithetischen Darstellung eines Problems. Aus den antithetischen Darstellungen lassen sich ebenfalls die konträr zueinander stehenden Annahmen bezüglich der theoretischen Natur des Problems ableiten. Die Darstellungsmethodik folgt dabei dem rational-deduktiven Ansatz. Die beiden rational-deduktiv gewonnenen Darstellungen werden dann demselben empirisch-induktiven Ansatz unterworfen, um die Art der zugrundeliegenden Annahmen aufzuzeigen. Sie sollen zeigen, daß identische Datenmengen dazu verwendet werden können, um jegliches theoretische Modell zu stützen. Dieses hat seinen Grund in der Tatsache, daß Daten keine Informationen sind, sondern erst Informationen durch Interpretation der Daten gewonnen werden, und zwar durch dialektische Konfrontation konträr zueinander stehender Interpretationen. Nach Ansicht von MASON und MITROFF ist dieser Ansatz für schlechtstrukturierte Entscheidungsprobleme geeignet, bei denen eine Debatte über die tatsächliche Natur dieses Entscheidungsproblems stattfindet.

Der kybernetisch-lernorientierte Ansatz basiert auf regelungstheoretischem Gedankengut, das durch Gesichtspunkte der Messungstheorie¹ erweitert wird. Dieser Ansatz geht davon aus, daß durch Rückkoppelungen Lernprozesse bei den Entscheidungsträgern induziert werden, die es gestatten, die Problemstruktur dergestalt zu modifizieren, daß schlecht-strukturierte in wohl-strukturierte Entscheidungen transformiert werden können und umgekehrt. Nach Auffassung von MASON und MITROFF² ist dieser Ansatz für alle Problemklassen anwendbar, wodurch er für die wissenschaftstheoretische Vorgehensweise beim Aufbau von Informationssystemen besonders geeignet erscheint.

Die letzte Komponente der Informationssystem-Definition betrifft die Darstellung der Information. Da der Marketing-Manager Probleme immer als begriffliche Gesamtheit erfaßt und eine Abbildung eines Problems im Lauf der Problemlösung aufbaut, wird er Daten zu allgemeinen Aussagen über das Problem reduzieren und so in eine begriffliche Gesamtheit umsetzen. Darüber hinaus überführt er die Daten in sein bereits existierendes internes Umweltmodell, wodurch dieses dadurch bestätigt oder ggf. modifiziert wird. Der Marketing-Manager urteilt über diese begriffliche Gesamtheit, die sich ihm aus den Daten anbietet, und diese werden zum Bestandteil seines internen Umweltmodells.³ Wegen der notwendigen Reduktion von Daten auf allgemeine Aussagen ist der Marketing-Manager geneigt, eher jene Informationsquellen zu verwenden, die ihm den intellektuellen Umsetzungsprozeß von Daten in Aussagen so ersparen. So wird er z.B. direkten Anschauungen und Meinungen über einen Tatbestand den Vorrang geben, wenn Daten wegen ihres Sur-

1 Vgl. dazu Churchman, C.W., Prediction and Optimal Decision: Philosophical Issues of a Science of Values, Englewood Cliffs, N.J. 1961, S. 339ff.

2 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 482.

3 Vgl. Le Breton, P.P. / Henning, D.A., Planning Theory, Englewood Cliffs, J.J. 1961, S. 260.

rogatcharakters die ihn interessierenden Aspekte nur höchst unvollkommen wiedergeben. Diese Feststellung impliziert, daß die Darstellungsform der Information den Bedürfnissen und dem psychologischen Typ des Marketing-Managers entsprechen muß. Es lassen sich in Anlehnung an MASON und MITROFF zwei Thesen aufstellen, die über künftig mögliche Forschungsrichtungen auf dem Gebiet der Informationssysteme Auskunft geben:¹

Die erste These lautet, daß der Manager solche Informationen benötigt, die auf seine persönliche Struktur zugeschnitten sind und nicht auf die der Designer. Dieses führt zu einer schweren Aufgabe für den Designer. Er muß nicht nur herausfinden, welche Informationen der Manager tatsächlich benötigt, sondern auch welche Darstellungsform (Computerausdrücke, Tabellen etc.) bezüglich des psychologischen Typs dieses Managers geeignet sind.

Die zweite These lautet, daß der Manager eine Methode der Erkenntnisgewinnung benötigt, die auf seine Probleme und nicht auf die der anderen an der Entwicklung eines Informationssystems Beteiligten zugeschnitten ist.

Das Resultat der Analyse von MASON und MITROFF² besteht darin, daß die festgestellten Zustände der fünf als variabel angesehenen Komponenten der Definition eines Informationssystems miteinander kombiniert werden können, um mögliche Entwicklungsrichtungen aufzuzeigen. Unter Vernachlässigung sehr spezifischer Ausprägungen von Variablen ergibt sich folgende Übersicht:

1. Psychologischer Typ des Benutzers
 - a. rational-sensitiver Typ
 - b. rational-intuitiver Typ

1 Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 485.

2 Ebenda, S. 476. (Die Nomenklatur wurde vom Verfasser teilweise geändert und die Übersicht gekürzt)

- c. emotional-sensitiver Typ
- d. emotional-intuitiver Typ

2. Problemstrukturierung

- a. wohlstrukturierte Entscheidungen
- b. schlechtstrukturierte Entscheidungen

3. Methodik der Erkenntnisgewinnung

- a. empirisch-induktiver Ansatz
- b. rational-deduktiver Ansatz
- c. kritizistisch-subjektiv-idealistischer Ansatz
- d. dialektisch-objektiv-idealistischer Ansatz
- e. kybernetisch-lernorientierter Ansatz

4. Problemklassen in der Organisation

- a. Strategische Planung
- b. Probleme auf Management-Ebene
- c. Probleme auf operativer Ebene

5. Darstellungsform der Information

- a. persönlich orientierte Darstellungsform
- c. unpersönliche Darstellungsform

Werden die Merkmalsausprägungen der fünf für ein Informationssystem konstitutiven Variablen miteinander kombiniert, ergeben sich $4 \times 2 \times 5 \times 3 \times 2 = 240$ Möglichkeiten, ein Informationssystem aufzubauen.¹

Ohne hier auf die Fülle möglicher Vorgehensweisen im Detail einzugehen, kann festgestellt werden, daß das Schwergewicht der gegenwärtigen Entwicklung auf dem Gebiet der Marketing-Informationssysteme beim denkbar-sensitiven Benutzer liegt, der wohlstrukturierte Probleme auf operativer Ebene mittels empirisch-induktiver Vorgehensweise löst und Informationen in unpersönlicher Darstellungsform erhält.

¹ Vgl. Mason, R.O. / Mitroff, I.I., A Program for Research on Management Information Systems, a.a.O., S. 476.

Als Antithese läßt sich ein Informationssystem postulieren, das einem empfindend-intuitiven Entscheidungsträger bei schlecht-strukturierten Entscheidungsproblemen wie z.B. der strategischen Marketing-Planung Informationen liefert, die unter Beachtung des kybernetisch-lernorientierten Ansatzes gewonnen werden und welche in einer persönlich orientierten Weise dargestellt werden. Zwischen diesen beiden "Grenzwerten" wird sich die zukünftige Forschung auf dem Gebiet der Marketing-Informationssysteme aller Wahrscheinlichkeit nach zu bewegen haben.

Abschließend seien der Untersuchung einige Entwicklungstendenzen angefügt, die mit großer Wahrscheinlichkeit die weitere Entwicklung bestimmen werden:

Die Einführung computergestützter Informationssysteme für den Bereich programmierbarer Entscheidungen wird sich verstärken. Mit dieser Feststellung soll keineswegs unkritisch der These SCHELSKYs¹ zugestimmt werden, "daß die technischen Möglichkeiten ihre Anwendung erzwingen", denn "keine Technologie enthebt der Notwendigkeit, vor ihrem Einsatz nach dessen Ziel zu fragen."² Über die Ziele, die mittels eines Marketing-Informationssystems erreicht werden sollen, herrscht aber momentan keineswegs Konsens.

Die Bewältigung von Planungs- und Entscheidungsproblemen im Marketing ist nur bedingt mittels eines Informationssystems möglich. Bestenfalls wird es hier nur eine Computerunterstützung geben, die auf Routineaufgaben der Planung und Entscheidung beschränkt ist. Hierfür sind die den Planungs- und Entscheidungsproblemen des Marketing immanenten kreativen Komponenten verantwortlich zu machen, aber

1 Schelsky, H., Der Mensch in der wissenschaftlichen Zivilisation, in: Schelsky, H., Hrsg., Auf der Suche nach der Wirklichkeit, Köln 1965, S. 439ff, hier S. 439.

2 Harnischfeger, H., Planung in der Sozialstaatlichen Demokratie, a.a.O., S. 49.

auch organisatorische und politische Zwänge dürften einer verstärkten entscheidungsorientierten Ausrichtung von Marketing-Informationssystemen entgegenstehen.

Marketing-Informationssysteme werden primär bei nicht-innovatorischen Marketing-Problemstellungen einsetzbar sein, die im Rahmen vorgegebener Randbedingungen gelöst werden können, ohne daß sie verändert werden müssen. Diese Feststellung beruht auf der Tatsache, daß das Marketing-Management gegenüber angeblichen Innovationen berechtigt skeptisch sein muß, und sie darüber hinaus oft unberechtigter Weise zu verhindern weiß: "[Politik] ist an der Beseitigung von Dysfunktionalitäten und an der Vermeidung von systemgefährdenden Risiken, also nicht an der Verwirklichung praktischer Ziele, sondern an der Lösung technischer Fragen orientiert."¹ Die Lösung innovativer Probleme bedarf weniger einer umfangreichen Datenbasis als vielmehr eines allgemeinen Konsens über die Richtung, in welcher die zu suchenden Problemlösungen gefunden werden können.

Echte Rationalisierungseffekte werden sich durch den Einsatz computergestützter Marketing-Informationssysteme nur als Nebeneffekte - etwa in der Form der Freisetzung von zeitaufwendigen Routinearbeiten - ergeben. Es ist momentan nicht absehbar, ob gegenüber den sehr komplexen Phänomenen im Marketing der jeweils immer gewonnene Rationalisierungsgrad durch gesteigerte Komplexitätswahrnehmung bei den zu lösenden Problemen überkompensiert wird oder nicht. Je komplexer insbesondere die Entscheidungssituationen werden, um so mehr besteht oft beim Marketing-Management die Neigung, den "status quo" aufrechtzuerhalten. Der Engpaß bei der Einführung und Entwicklung von Marketing-Informationssystemen wird somit oft durch mangelnde Innovationsbereitschaft des Management bestimmt. Dieser

¹ Habermas, J., Technik und Wissenschaft als "Ideologie", in: Habermas, J., Hrsg., Technik und Wissenschaft als "Ideologie", 5. Aufl., Frankfurt 1971, S. 48ff, hier S. 77.

Engpaß kann nur geringfügig durch Marketing-Informationssysteme beseitigt werden. Es bedarf darüber hinaus innovativer und damit bezüglich der Realisierung dringend notwendiger organisatorischer Ansätze.

Literaturverzeichnis

- Ackoff, R.L., Scientific Method - optimizing applied research decisions, New York-London 1962
- Adam, A., Messen und Regeln in der Betriebswirtschaft - Einführung in die informationswissenschaftlichen Grundzüge der industriellen Unternehmensforschung, Würzburg 1959
- Adam, A./
Fersch, F./
Klamecker, A. et al., Anwendung der Matrizenrechnung auf wirtschaftliche und statistische Probleme, 2. Aufl., Würzburg 1963
- Aiserman, M.A. et al., Logik - Automaten - Algorithmen, München-Wien 1967
- Albach, H., Wirtschaftlichkeitsrechnung bei unsicheren Erwartungen, Köln und Opladen 1959
- Albert, H., Probleme der Theoriebildung, in: Albert, H., Hrsg., Theorie und Realität, Tübingen 1964, S. 3ff
- Anderson, O., Probleme der statistischen Methodenlehre in den Sozialwissenschaften, 3. Aufl., Würzburg 1957
- Ansoff, H.I., A Quasi-Analytic Method for Long Range Planning, in: Alexis, M./ Wilson, C.Z., Hrsg., Organizational Decision Making, Englewood Cliffs, N.J. 1967, S. 427ff
- Bachman, C.W., Integrated Data Store, in: DPMA Quarterly, January 1965, S. 10ff

- Bachman, C.W./
Williams, S.B., A General Purpose Programming System for Random Access Memories, in: Proceedings of the Fall Joint Computer Conference, Part I, San Francisco, Oktober 1964, S. 411ff
- Baetge, J./
Steenken, H.-U., Theoretische Grundlagen eines Regelungsmodells zur operationalen Planung und Überwachung betriebswirtschaftlicher Prozesse, in: Zeitschrift für betriebswirtschaftliche Forschung, 23. Jg., 1971, S. 593ff
- Baetge, J./
Steenken, H.-U., Regelungstheoretischer Ansatz zur operationalen Planung und Überwachung von Produktion und Lagerhaltung, in: Zeitschrift für betriebswirtschaftliche Forschung, 24. Jg., 1972, S. 22ff
- Barnard, C., The Functions of the Executive, Cambridge 1938
- Barr, J.J., SIC: a basic tool for the marketer, in: Industrial Marketing, August 1969, S. 52f und S. 76ff
- Basler, H., Grundbegriffe der Wahrscheinlichkeitsrechnung und statistische Methodenlehre, Würzburg-Wien 1968
- Bates, W.S., Security of Computer Based Information Systems, in: Datamation, May 1970, S. 60ff
- Behnke, H. et al., Mathematik, Bd. 1, Frankfurt 1964
- Behnke, H. et al., Mathematik, Bd. 2, Frankfurt 1966
- Behrens, K.Ch., Marktforschung, 2. Aufl., Wiesbaden 1966

- Benner, F.H., On designing generalized file records for management information systems, in: AFIPS Conference Proceedings, Vol. 31, 1967, S. 291ff
- Berge, C., The Theory of Graphs and its Applications, London-New York 1964
- Bertalanffy, L., General Systems Theory, in: Human Biology, 1951, S. 303ff
- Bidlingmaier, J., Die Bedeutung zielgerichteter Informationen für die Unternehmensführung, in: Zeitschrift für Organisation, 39. Jg., 1970, S. 97ff
- Bierfelder, W., Optimales Informationsverhalten im Entscheidungsprozeß der Unternehmung, Berlin 1968
- Biermann, H., Kybernetische Prognosemodelle in der Regionalforschung, Bd. 2 der Schriftenreihe "Wirtschaftskybernetik und Systemanalyse", hrsg. v. K.-E. Schenk, Berlin 1970
- Blau, P.M., Die Dynamik bürokratischer Strukturen, in: Mayntz, R., Hrsg., Bürokratische Organisation, 2. Aufl., Köln-Berlin 1971, S. 310ff
- Blumenthal, S.C., Management Information Systems - A Framework for Planning and Development, Englewood Cliffs, N.J. 1969
- Bochenski, I.M., Die zeitgenössischen Denkmethoden, 3. Aufl., Bern-München 1965
- Bock, R.D., Contributions of Multivariate Experimental Designs to Educational Research, in: Cattell, R.B., Hrsg., Handbook of Multivariate Experimental Psychology, Chicago 1966, S. 820ff

- Borrow, D.G./
Raphael, B.,
A Comparison of List-Processing Languages, in: Communications of the Association for Computing Machinery, New York 1964, S. 231ff
- Bössmann, E.,
Die ökonomische Analyse von Kommunikationsbeziehungen in Organisationen, Berlin usw. 1967
- Bräutigam, J./
Küllmer, H.,
Die Anwendung statistischer Verfahren zur Objektivierung der Kreditwürdigkeitsprüfung, in: Betriebswirtschaftliche Blätter, 21. Jg., 1972, S. 6ff
- Buchholz, W.,
File Organization and Adressing, in: IBM Systems Journal, Vol. 2, 1963, S. 86ff
- Bull-General
Electric,
IDS Einführung, Ref.-Nr. 18.12.080 D, März 1968
- Busacker, R.G./
Saaty, Th.L.,
Endliche Graphen und Netzwerke, München-Wien 1968
- Carnap, R./
Stegmüller, W.,
Induktive Logik und Wahrscheinlichkeit, Wien 1959
- Cattell, R.B.,
The Data Box: Its Ordering Of Total Ressources In Term Of Possible Relational Systems, in: Cattell, R.B., Hrg., Handbook of Multivariate Experimental Psychology, Chicago 1966, S. 67ff
- Chmielewicz, K.,
Forschungskonzeptionen der Wirtschaftswissenschaft - Zur Problematik einer entscheidungstheoretischen und normativen Wirtschaftslehre, Stuttgart 1970

- Churchman, C.W., Prediction and Optimal Decision: Philosophical Issues of a Science of Values, Englewood Cliffs, N.J. 1961
- Churchman, C.W., The Design of Inquiring Systems, New York 1971
- Claycamp, H.J./
Amstutz, A.E., Simulation Techniques in Analysis of Marketing Strategy, in: Bass, F.M./King, C.W./Pessemier, E.A., Hrsg., Applications of the Sciences in Marketing Management, New York usw. 1968, S. 113ff
- Cleland, D./
King, W.R., Systems Analysis and Project Management, New York 1968
- Cloonan, J.B., A Heuristic Approach to Some Sales Territory Problems, in: Proceedings Of the Fourth International Conference On Operational Research, hrsg. v. D.B. Hertz/J. Malese, New York usw. 1966, S. 284ff
- Cochran, W.G., The Design of Experiments, in: Flagle, C.D./Huggins, W.H./Roy, R.H., Operations Research and Systems Engineering, Baltimore 1960, S. 508ff
- CODASYL, Data Base Task Group Report, New York 1971
- CODASYL Systems
Committee, Feature Analysis Of Generalized Data Base Management Systems, New York 1971
- Cooley, W.W./
Lohnes, P.R., Multivariate Data Analysis, New York 1971
- Craik, K., The Nature of Explanation, Cambridge 1943

- Cyert, R./
March, J.,
The Behavioral Theory of the Firm - A Behavioral Science-Economics Amalgam, in: Cooper, W./Leavitt, H./Shelly M., Hrsg., New Perspectives in Organization Research, New York 1964, S. 289ff
- Deming, W.E.,
Some Theory of Sampling, New York 1950
- Dichtl, E.,
Über Wesen und Struktur absatzpolitischer Entscheidungen, Berlin 1967
- Dreger, W.,
Management-Informationssysteme, Wiesbaden 1973
- Dworatschek, S.,
Einführung in die Datenverarbeitung, Berlin-New York 1971
- Emery, J.C.,
Organizational Planning and Control Systems - Theory and Technology, New York-London 1969
- Enger, N.L.,
Putting MIS to Work - Managing the Management Information System, New York 1969
- Ernst, G.W.,
GPS and Decision Making - An Overview, in: Banerji, R.B./Mesarovic, M.D., Hrsg., Theoretical Approaches to Non-Numerical Problem Solving, Heidelberg-New York 1970, S. 59ff
- Fäßler, K.,
Betriebliche Mitbestimmung - Verhaltenswissenschaftliche Projektionsmodelle, Bd. 6 der Schriftenreihe "Die Betriebswirtschaft in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1970
- Fellner, W.,
Probability and Profit - A Study of Economic Behavior along Bayesian Lines, Homewood, Ill. 1965

- Feltham, G.A., The Value of Information, in: Accounting Review, Vol. XLIII, 1968, S. 684ff
- Ferguson, G.A., Statistical Analysis in Psychology and Education, 2. Aufl., New York usw. 1966
- Fisher, R.A., The use of multiple measurements in taxonomic problems, in: Annales of Eugenics, Vol. 7, 1939, S. 179ff
- Fisher, R.A., Statistische Methoden für die Wissenschaft, 12. Aufl., Edinburgh-London 1956
- Fisk, G., Marketing-Systems - An Introductory Analysis, New York usw. 1967
- Flechtner, H.-J., Grundbegriffe der Kybernetik - Eine Einführung, Stuttgart 1969
- Fletcher, A./
Clarke, G., Mathematische Hilfsmittel der Unternehmensführung, München 1966
- Forrester, J.W., Industrial Dynamics, 2. Aufl., Cambridge, Mass. 1962
- Förster, E./
Egermayer, F., Korrelations- und Regressionsanalyse - Ein Leitfaden für Ökonomen, Berlin 1966
- Frank, R.E./
Green, P.E., Numerical Taxonomy in Marketing Research - A Review Article, in: Journal of Marketing Research, February 1968, S. 83ff
- Frank, R.E./
Green, P.E., Quantitative Methods in Marketing, Englewood Cliffs, N.J. 1967

- Freter, H., Mediaselektion - Informationsgewinnung und Entscheidungsmodelle für die Werbeträgerauswahl, Bd. 4 der Schriftenreihe "Unternehmensführung und Marketing", Hrsg. H. Meffert, Wiesbaden 1974 (im Druck)
- Gatty, R., Multivariate Analysis for Marketing Research - An Evaluation, in: Applied Statistics, Vol. 15, 1960, S. 157ff
- Gäfigen, G., Theorie der wirtschaftlichen Entscheidung - Untersuchungen zur Logik und ökonomischen Bedeutung des rationalen Handelns, Tübingen 1963
- Gesellschaft für Mathematik und Datenverarbeitung m.b.H. Bonn (GMD) Abteilung behördliche DV Systeme, Kurzbeschreibung von Information Storage and Retrieval Systemen, Heft 2, Schloss Birlinghoven, September 1972
- Gödel, K., Über formal unterscheidbare Sätze der Principia Mathematica und verwandte Systeme, in: Monatshefte für Mathematik und Physik, 38. Jg., 1931, S. 173ff
- Grochla, E., Die Integration der Datenverarbeitung, in: Bürotechnik und Automation, 9. Jg. 1968, S. 108ff
- Grochla, E., Forschung und Entwicklung auf dem Gebiet der Informationssysteme als Aufgabe der Betriebswirtschaftslehre, in: Zeitschrift für Betriebswirtschaft, 41. Jg., 1971, S. 563ff
- Grochla, E., Modelle und betriebliche Informationssysteme, Arbeitspapier No. 1, Februar 1973
- de Grolier, E., A Study of general categories applicable to classification and coding in documentation, Paris 1962

- Gutenberg, E., Grundlagen der Betriebswirtschaftslehre, Erster Band - Die Produktion, 11. Aufl., Berlin-Heidelberg-New York 1965
- Habermas, J., Technik und Wissenschaft als "Ideologie", in: Habermas, J., Hrsg., Technik und Wissenschaft als "Ideologie", 5. Aufl., Frankfurt 1971, S. 48ff
- Harnischfeger, H., Planung in der Sozialstaatlichen Demokratie, Berlin 1969
- Hax, H., Die Koordination von Entscheidungen, Köln und Opladen 1965
- Heinen, E., Das Zielsystem der Unternehmung - Grundlagen betriebswirtschaftlicher Entscheidungen, Wiesbaden 1966
- Henry, W.R., Hierarchical Structure for Data Management, in: IBM Systems Journal, Vol. 8, No. 1, 1969, S. 2ff
- Hermes, H., Aufzählbarkeit - Entscheidbarkeit - Berechenbarkeit, Berlin-Göttingen-Heidelberg 1961
- Hilbert, D./ Ackermann, W., Grundzüge der theoretischen Logik, 5. Aufl., Berlin usw. 1967
- Hirsch, R.E., Informationswert und -kosten und deren Beeinflussung, in: Zeitschrift für betriebswirtschaftliche Forschung, 20. Jg., 1968, S. 670ff
- Hofstätter, P., Psychologie, Frankfurt 1957
- Holbrook, J.G., Laplace-Transformationen, Braunschweig 1970

- Hooke, R./
Jeeves, T.A., Comment on Brook's Discussion of
Random Methods, in: Operations
Research, Vol. 6, 1958, S. 881f
- Hosse, H., Landesdatenbank Nordrhein-West-
falen, in: ADL-Nachrichten, 13.
Jg., 1968, S. 320ff
- Hummel, F.E., Market and Sales Potentials, New
York 1961
- Hüttner, M., Grundzüge der Marktforschung,
Wiesbaden 1972
- IBM, Datenfernverarbeitung - IBM Tele-
processing, Formular Nr. 74968-1,
Juli 1966
- IBM, Datenorganisation auf Speicherein-
heiten mit direktem Zugriff, IBM-
Formular Nr. 79966-1, o.O. (Sindelf-
ingen) 1968
- IBM, Ein Weg zu einem Führungsinformati-
onssystem, Management-Einführungs-
schrift, September 1970
- IBM, Customer Information Control Systems
Application Description, Formular
Nr. CH 20-0318-6, New York 1970
- IBM, Introduction to List Processing
Facilities of PL/1, IBM-Formular
Nr. 20-0015
- Isler, M., Selecting Data for Community Rene-
val in: Journal of American Insti-
tute of Planners, March 1967, S. 74ff
- Jahn, W./
Vahle, H., Die Faktorenanalyse und ihre Anwendung,
Berlin 1970

- Jones, L.V., Analysis of Variance in its Multi-
variate Developments, in: Cattell,
R.B., Hrsg., Handbook of Multivariate
Experimental Psychology, Chicago
1966, S. 244ff
- Kamlah, W./ Logische Propädeutik oder Vorschule
Lorenzen, P., des vernünftigen Redens, Mannheim
1967
- Kappler, E., Systementwicklung - Lernprozesse in
betriebswirtschaftlichen Organisatio-
nen, Bd. 11 der Schriftenreihe "Die
Betriebswirtschaftslehre in Forschung
und Praxis", Hrsg. E. Heinen, Wies-
baden 1972
- Katterle, S., Normative und explikative Betriebs-
wirtschaftslehre, Göttingen 1964
- Katz, D./ The Social Psychology of Organizations,
Kahn, R.L., New York usw. 1967
- Kellerer, H., Theorie und Technik des Stichproben-
verfahrens - Eine Einführung unter be-
sonderer Berücksichtigung der Anwendung
auf soziale und wirtschaftliche Massen-
erscheinungen, Einzelschriften der
Deutschen Statistischen Gesellschaft,
Nr. 5, 3. Aufl., München 1963
- Kirsch, W., Entscheidungsprozesse, Bd. II: In-
formationsverarbeitungstheorie des
Entscheidungsverhaltens, Wiesbaden
1971
- Kirsch, W., Entscheidungsprozesse, Bd. III: Ent-
scheidungen in Organisationen, Wies-
baden 1971
- Kirsch, W./ Organisationstheorien und Betriebs-
Meffert, H., wirtschaftslehre, Bd. 1 der Schriften-
reihe der Zeitschrift für Betriebs-
wirtschaft, Hrsg. E. Gutenberg, Wies-
baden 1970

- Klaus, G., Hrsg., Wörterbuch der Kybernetik, 2. Aufl., Berlin 1968
- Kleene, S.C., Introduction to Metamathematics, 2. Aufl., Amsterdam 1964
- Klein, H., Heuristische Entscheidungsmodelle - Neue Techniken des Programmierens und Entscheidens für das Management, Bd. 9 der Schriftenreihe "Die Betriebswirtschaft in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1971
- Kneale, W., Probability and Induction, Oxford 1949
- Knight, F., Risk, Uncertainty and Profit, 5. Aufl., London 1940
- Koller, S./
Schaeffer, K.A., Kap.: Beurteilung der Genauigkeit, in: Stichproben in der amtlichen Statistik, hrsg. vom Statistischen Bundesamt, Wiesbaden 1960, S. 49ff und S. 97ff
- Kosiol, E., Modellanalyse als Grundlage unternehmerischer Entscheidungen, in: Zeitschrift für handelswissenschaftliche Forschung, 13. Jg., 1961, S. 318ff
- Kosiol, E., Die Unternehmung als wirtschaftliches Aktionszentrum - Einführung in die Betriebswirtschaftslehre, Reinbek 1966
- Kotler, P., Die Zukunft des Computers im Marketing, in: Der Markt, 1970, S. 81ff
- Koxholt, R., Die Simulation - Ein Hilfsmittel der Unternehmensforschung, München-Wien 1967

- Köhler, R., Informationssysteme für die Unternehmensführung, in: Zeitschrift für Betriebswirtschaft, 41. Jg., 1971, S. 27ff
- Kramer, R., Information und Kommunikation - Betriebswirtschaftliche Bedeutung und Einordnung in die Organisation der Unternehmung, Betriebswirtschaftliche Forschungsergebnisse, Bd. 23, Hrsg. E. Grochla, Berlin 1965
- Krautter, J., Marketing-Entscheidungsmodelle, Bd. 11 der "Schriften zur theoretischen und angewandten Betriebswirtschaftslehre", Hrsg. L. Pack, Wiesbaden 1973
- Krelle, W., Preistheorie, Tübingen-Zürich 1968
- Krelle, W., Präferenz- und Entscheidungstheorie, Tübingen 1968
- Kuehn, A.A./
Hamburger, M.J., A heuristic program for locating warehouses, in: Management Science, Vol. 9, 1962/63, S. B-643ff
- Kutschker, M., Verhandlungen als Elemente eines verhaltenswissenschaftlichen Bezugsrahmens des Investitionsgütermarketing, Diss., Mannheim 1972
- Landgraf, C./
Schneider, G., Elemente der Regelungstechnik, Berlin-Heidelberg-New York 1970
- Le Breton, P.P./
Henning, D.A., Planning Theory, Englewood Cliffs, N.J. 1961
- Leffson, U., Die Grundsätze ordnungsmäßiger Buchführung, 3. Aufl., Düsseldorf 1972

- Lenk, H., Läßt sich die Erkenntnis kybernetisch begründen?, in: Göttingische gelehrte Anzeigen, 219. Jg., 1967, S. 131ff
- Lenz, H., Grundlagen der Elementarmathematik, Berlin 1960
- Lewandowski, R., Die Problematik der kurzfristigen Vorhersage und ihre Lösung mit dem Programm-System PROCOM, Manuskript des Mathematischen Beratungs- und Programmierdienst G.m.b.H., Dortmund o.J.
- Lindemann, P., Die Organisation des Rechnungswesens bei automatisierter Datenverarbeitung, in: IBM Nachrichten, 18. Jg., 1968, S. 177ff
- Linder, A., Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure, 2. Aufl., Basel-Stuttgart 1957
- Little, J.D.C., Models and Managers: The Concept of a Decision Calculus, in: Management Science, Vol. 16, 1970, S. B-466ff
- Loitlsberger, E., Zum Informationsbegriff und zur Frage der Auswahlkriterien von Informationsprozessen, in: Loitlsberger, E., Hrsg. Empirische Betriebswirtschaftslehre, Festschrift zum 60. Geburtstag von L.L. Illetschko, Wiesbaden 1963, S. 115ff
- Ludwig, R., Methoden der Fehler- und Ausgleichsrechnung, Braunschweig 1969
- Luhmann, N., Politische Planung, in: Jahrbuch für Sozialwissenschaften, 17. Jg., Bd. 1, 1966, S. 271ff

- Luhmann, N., Zweckbegriff und Systemrationalität - Über die Funktion von Zwecken in sozialen Systemen, Tübingen 1968
- Lutz, Th., Was ist eine Datenbank?, in: Bürotechnik und Automation, 8. Jg., 1967, S. 250ff
- Lutz, Th., Direktanfragen an eine Datenbank, in: Bürotechnik und Automation, 9. Jg., 1968, S. 96ff
- Lutz, Th., Adressketten und Datentechnik, in: Bürotechnik und Automation, 9. Jg., 1968, S. 178ff
- Lutz, Th., Generalized Information System (GIS) - Ein Exekutivsystem für Datenbankprobleme, Teil 2: Die GIS-Sprache, in IBM Nachrichten, 19. Jg., 1969, S. 714ff
- Lutz, Th., Formatierte Datenbanksätze, in: Bürotechnik und Automation, 10. Jg., 1969, S. 48ff
- Lutz, Th., Formale Aspekte des Informationssystems und Schlußfolgerungen für ein MIS, in: IBM Nachrichten, 22. Jg., 1972, S. 54ff
- Lutz, Th./
Beutler, H., Management Information Systems (MIS) - Begriffe und Konzeptionen für Management-Informationssysteme, in: IBM Nachrichten, 18. Jg., 1968, S. 367ff
- Lutz, Th./
Klimesch, H., Management Information Systems (MIS) II - Die Datenbank und ihre Probleme, in: IBM Nachrichten, 18. Jg., 1968, S. 457ff
- Lutz, Th./
Klimesch, H., Die Datenbank im Informationssystem, München-Wien 1971

- Mach, E., Erkenntnis und Irrtum, 4. Aufl.,
Leipzig 1920
- McMillan, C./ Systems Analysis - A Computer Approach
Gonzales, R.F., to Decision Models, Homewood, Ill.
1965
- v. Mangoldt, H./ Einführung in die höhere Mathema-
Knopp, K., tik, Bd. 3, 9. Aufl., Stuttgart 1948
- Marschak, J., Towards an Economic Theory of Orga-
nization and Information, in: Thrall,
R.M./Coombs, C.H./Davis, R.L., Hrsg.,
Decision Processes, New York-London
1954, S. 184ff
- Mason, R.O./ A Program for Research on Management
Mitroff, I.I., Information Systems, in: Management
Science, Vol. 19, 1973, S. B-475ff
- Massy, W.F., Discriminant Analysis of Audience
Characteristics, in: Journal of Ad-
vertising Research, Vol. 5, 1965,
S. 39ff
- Meadow, Ch.T., The Analysis of Information Systems -
A Programmer's Introduction to In-
formation Retrieval, New York-London-
Sydney 1967
- Meffert, H., Betriebswirtschaftliche Kosteninfor-
mationen - Ein Beitrag zur Theorie der
Kostenrechnung, Bd. 4 der Schriften-
reihe "Die Betriebswirtschaft in For-
schung und Praxis", Hrsg. E. Heinen,
Wiesbaden 1968
- Meffert, H., Die Flexibilität in betriebswirt-
schaftlichen Entscheidungen, Habili-
tationsschrift, München 1968 (unver-
öffentlicht)
- Meffert, H., Zum Problem der betriebswirtschaft-
lichen Flexibilität, in: Zeitschrift
für Betriebswirtschaft, 39. Jg., 1969,
S. 779ff

- Meffert, H., Art.: Marketing, in: Management-Enzyklopädie, Bd. IV, München 1971, S. 383ff
- Meffert, H., Modelle des Käuferverhaltens und ihr Aussagewert für das Marketing, in: Zeitschrift für die gesamten Staatswissenschaften, Bd. 127, 1971, S. 327ff
- Meffert, H., Unternehmensziele, in: Schöttle, K.M., Hrsg., Jahrbuch des Marketing, Essen 1971, S. 22ff
- Meffert, H., Die Leistungsfähigkeit der entscheidungs- und systemorientierten Marketing-Theorie, in: Wissenschaftsprogramm und Ausbildungsziele der Betriebswirtschaftslehre, Tagungsberichte des Verbandes der Hochschullehrer für Betriebswirtschaft e.V., Bd. I, hrsg. v. G. v. Kortzfleisch, Berlin 1971, S. 167ff
- Meffert, H., Systemtheorie aus betriebswirtschaftlicher Sicht, in: Wirtschaftskybernetik und Systemanalyse, Bd. 1, hrsg. v. K.-E. Schenk, Berlin 1971, S. 174ff
- Meffert, H., Informationssysteme - EDV und Systemanalyse, Teil I: Grundlagen, Münster 1973
- Meffert, H., Gute Zukunftschancen für mathematische Modelle, in: Wirtschaftswoche, 27. Jg., Nr. 19, 1973, S. 75ff
- Meffert, H., Art.: Absatztheorie, systemorientierte, in: Tietz, B., Hrsg., Handwörterbuch der Absatzwirtschaft (im Druck)

- Meffert, H., Computergestützte Marketing-Informationssysteme und Marketing-Modelle: Gegenwärtiger Stand und Entwicklungstendenzen, unveröffentlichtes Manuskript eines Beitrags zum Wirtschaftsinformatiksymposium 1973 der IBM Deutschland über computergestützte Marketing-Planung
- Meffert, H. und Mitarbeiter, Die Anwendung mathematischer Modelle im Marketing, in: Jacob, H., Hrg., Schriften zur Unternehmensführung, Bd. 14, S. 92ff und Bd. 15, S. 23ff
- Minsky, M., Steps Toward Artificial Intelligence, in: Feigenbaum, E.A./Feldman, J., Hrg., Computers and Thought, New York 1963, S. 406ff
- Moloney, R.F., New Generation EDP Control Considerations, in: Management Services, March/April 1968, S. 17ff
- Montgomery, D.B., Marketing Information and Decision Systems: Coming of Age in the '70's, Graduate School of Business - Stanford University, Research Paper No. 150, April 1973
- Montgomery, D.B./Urban, G.L., Marketing Decision-Information Systems: An Emerging View, in: Journal of Marketing Research, Vol. VII, 1970, S. 226ff
- Morgenstern, O., Über die Genauigkeit wirtschaftlicher Beobachtungen, 2. Aufl., Wien-Würzburg 1965
- Morris, W.T., Intuition and Relevance, in: Management Science, Vol. 14, 1967, S. B-157ff
- Müller, W., Technik und Leistungsfähigkeit betriebswirtschaftlicher Simulationsstudien, in: Zeitschrift für Betriebswirtschaft, 38. Jg., 1968, S. 605ff

- Naschhold, F., Systemsteuerung, Stuttgart 1969
- Naylor, T.H., Marketing-Politik-Modelle, o.O.
und o.J.
- Newell, A./
Shaw, J.C./
Simon, H.A., The Process of Creative Thinking,
in: Gruber, H.E./Terrell, G./
Wertheimer, M., Hrsg., Contem-
porary Approaches to Creative
Thinking, New York 1962, S. 63ff
- Newell, A./
Shaw, J.C./
Simon, H.A., Chess-Playing Programs and the
Problem of Complexity, in:
Feigenbaum, E.A./Feldman, J.,
Computers and Thought, New York
1963, S. 39ff
- Newell, A./
Tonge, F.M., An Introduction to Information
Processing Language-V, in:
Communications of the Association
for Computing Machinery, New
York 1960, S. 205ff
- Nicosia, F.M., Consumer Decision Processes,
Englewood Cliffs, N.J. 1966
- Niedereichholz, J., Betriebswirtschaftslehre und
Informatik, in: Wissenschafts-
programm und Ausbildungsziele
der Betriebswirtschaftslehre,
Tagungsberichte des Verbandes der
Hochschullehrer für Betriebswirt-
schaft e.V., Bd. I, hrsg. v. G.
v. Kortzfleisch, Berlin 1971,
S. 225ff
- Nieschlag, R./
Dichtl, E./
Hörschgen, H., Marketing, 4. Aufl., Berlin 1971
- Osgood, C.E., The Nature and Measurement of
Meaning, in: Psychological Bulle-
tin, 49. Jg., 1952, S. 197ff

- Osgood, C.E./
Suci, G.J./
Tannenbaum, P.H.,
The Measurement of Meaning,
Urbana, Ill. 1957
- Oswald, A.,
Datenschutz - ein hessisches
Modell, in: IBM Nachrichten,
20. Jg., 1970, S. 379ff
- Pagel, K./
Eberhard, W.,
Management Informationssystem
MIS/360 - Erste Rationalisierungs-
stufe im Verwaltungsbereich der
IBM Deutschland, Teil I: An-
wendungen und Möglichkeiten des
MIS/360-Systems, in: IBM Nach-
richten, 20. Jg., 1970, S. 463ff
- Pagés, R.,
Das Experiment in der Soziologie,
in: König, R., Hrsg., Handbuch
der Empirischen Sozialforschung,
Stuttgart 1962, S. 415ff
- Perreault, J.M.,
Towards a theory for UDC - Essays
aimed at structural understanding
and operational improvement,
London 1969
- Piersol, R.J.,
Accuracy of Estimating Markets
for Industrial Products by Size
of Consuming Industries, in:
Journal of Marketing, Vol. V,
1968, S. 147ff
- Popper, K.R.,
Truth, Rationality, and the
Growth of Scientific Knowledge,
in: Popper, K.R., Hrsg., Con-
jectures and Refutations - The
Growth of Scientific Knowledge,
New York-London 1962, S. 242ff
- Popper, K.R.,
Das Elend des Historizismus,
Tübingen 1965
- Popper, K.R.,
Logik der Forschung, 3. Aufl.,
Tübingen 1969

- Reitman, W.R., Heuristic Decision Procedures, Open Constraints, and the Structure of Illdefined Problems, in: Shelly, M.W./Glenn, L.B., Hrsg., Human Judgements and Optimality, New York-London 1964, S. 282ff
- Reusch, G., Die Struktur des Programmsystems und der Datenbank für ein Informationssystem, in: ADL-Nachrichten, 13. Jg., 1968, S. 24ff
- Robertson, S./
Kennedy, J.N., Prediction of Consumer Innovators: An Application of Multiple Discriminant Analysis, in: Aaker, D.A., Hrsg., Multivariate Analysis in Marketing: Theory and Application, Belmont, Cal. 1971, S. 117ff
- Roßner, R., Varianzanalyse: Einfachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden I: Grundlagen und Versuchsplanung, hrsg. v. E. Walter, Berlin-Heidelberg-New York 1970, S. 217ff
- Roßner, R., Zweifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden I: Grundlagen und Versuchsplanung, hrsg. von E. Walter, Berlin-Heidelberg-New York 1970, S. 230ff
- Roßner, R., Dreifachklassifikation, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden und Versuchsplanung, hrsg. von E. Walter, Berlin-Heidelberg-New York 1970, S. 235ff
- Rummel, R.J., Understanding factor analysis, in: The Journal of Conflict Resolution - A Quarterly for Research Related to War and Peace, Vol. XI, No. 4, 1967, S. 444ff

- Russell, B., Human Knowledge, London 1951
- Rütschi, K., Marketing-Modelle - Systemanalyse: Formalisierung und Aufbau, Winterthur 1972
- Salomonsson, C., Experiences Of A Pilot Study On Data Banks For Development, in: Journées "Banque De Données", hrsg. von der Association Française pour la Cybernetique Economique et Technique et Institut de Recherche d'Information et d'Automatique, Aix-En-Provence, Juin 1971, S. 15.02ff
- Schauff, K., Datenbank-Management mit IDS, in: B-GE-Schriften, Heft 4, August 1970
- Schelski, H., Der Mensch in der wissenschaftlichen Zivilisation, in: Schelski, H., Hrsg., Auf der Suche nach der Wirklichkeit, Köln 1965, S. 439ff
- Schiemenz, B., Regelungstheoretische Grundkonzepte und ihre Anwendung zur Gestaltung von Produktions- und Lagerhaltungssystemen, in: Zeitschrift für betriebswirtschaftliche Forschung, 24. Jg., 1972, S. 434ff
- Schlaifer, R., Probability and Statistics for Business Decisions, New York 1959
- Schlick, M., Die Kausalität in der gegenwärtigen Physik, in: Naturwissenschaften, 19. Jg., 1931, S. 143ff
- Schmitt, H.J., Information und Strategie der Unternehmung in Marktkampfsituationen, in: Zeitschrift für Betriebswirtschaft, 37. Jg., 1967, S. 51ff

- Schmitt-Grohé, J., Produktinnovation - Verfahren und Organisation der Neuproduktplanung, Bd. 3 der Schriftenreihe "Unternehmensführung und Marketing", Hrsg. H. Meffert, Wiesbaden 1972
- Schneeweiß, H., Das Aggregationsproblem, in: Statistische Hefte, 6. Jg., 1965, S. 1ff
- Schröder, K., Mathematik für die Praxis - Ein Handbuch, Bd. I., Frankfurt-Zürich 1964
- Schulz, A., Strukturanalyse der maschinellen betrieblichen Informationsverarbeitung, Berlin 1970
- Schulz, R., Kaufentscheidungsprozesse des Konsumenten, Bd. 2 der Schriftenreihe "Unternehmensführung und Marketing", Hrsg. H. Meffert, Wiesbaden 1972
- Schulze, J.H., Datenschutz in der Datenverarbeitung, in: IBM Nachrichten, 21. Jg., 1971, S. 640ff
- Schweiker, K.F., Grundlageneiner Theorie betrieblicher Datenverarbeitung, Bd. 2 der Schriftenreihe des Betriebswirtschaftlichen Instituts für Organisation und Automation an der Universität zu Köln, Hrsg. E. Grochla, Wiesbaden 1966
- Seiffert, H., Einführung in die Wissenschaftstheorie - Sprachanalyse-Deduktion-Induktion in Natur- und Sozialwissenschaften, 3. Aufl., München 1971
- Shaw, J.C./
Newell, A./
Simon, H.A./
Ellis, T.O., A Command Structure for Complex Information Processing, in: Proceedings of the Western Joint Computer Conference, Los Angeles 1958, S. 119ff

- Sheth, J.N., Multivariate Analysis in Marketing,
in: Journal of Advertising Research,
Vol. 10, 1970, S. 29ff
- Sheth, J.N., The Multivariate Revolution in
Marketing Research, in: Journal of
Marketing, Vol. 35, 1971, S. 13ff
- Simon, H.A., Administrative Behavior - A Study
of Decision-Making Processes in
Administrative Organization, 2.
Aufl., New York 1957
- Spiegel, B., Die Struktur der Meinungsverteilung
im sozialen Feld, Bern-Stuttgart 1961
- Spies, P.P., Teilnehmer-Rechensysteme, in: Elek-
tronische Rechenanlagen, No. 9,
1967, S. 117ff
- Stachowiak, H., Über kausale, konditionale und
strukturelle Erklärungsmodelle,
in: Philosophia Naturalis, Bd. IV,
H. 4, 1957, S. 403ff
- Stachowiak, H., Denken und Erkennen im kybernetischen
Modell, 2. Aufl., Wien-New York 1969
- Steffenhagen, H., Konflikt und Koordination in Distri-
butionssystemen - Ansätze einer ver-
haltenswissenschaftlichen Gleichge-
wichtsanalyse, Diss., Münster 1972
- Stegmüller, W., Unvollständigkeit und Unentscheid-
barkeit, Hrsg., Wien 1959
- Stegmüller, W., Hauptströmungen der Gegenwarts-
philosophie - Eine kritische Ein-
führung, Stuttgart 1960
- Stern, M.E., Marketing Planung - Eine System-
Analyse, 2. Aufl. Berlin 1969

- Strecker, H., Nachprüfung als Mittel zur Feststellung systematischer Angabefehler in Erhebungen, in: Strecker, A./Bihn, W.R., Hrg., Die Statistik in der Wirtschaftsforschung, Festgabe für Rolf Wagenführ, Berlin 1967, S. 439ff
- Sussenguth, H.E., Use of Tree Structures for Processing Files, in: Communications of the Association for Computing Machinery, New York 1963, S. 272ff
- Süsswold, F., Bedingungen und Gesetzmäßigkeiten des Problemlösungsverhaltens, in: Graumann, C.F., Hrg., Denken, 4. Aufl., Berlin 1969, S. 273ff
- Szyperski, N., Gegenwärtiger Stand und Tendenzen der Entwicklung betrieblicher Informationssysteme, in: IBM Nachrichten, 23. Jg., 1973, S. 473ff
- Tarski, A., Undecible Theories, Amsterdam 1953
- Teichmann, H., Physikalische Anwendung der Vektor- und Tensorrechnungen, Mannheim 1964
- Tinter, E., Econometrics, New York 1952
- Ujomov, A.I., Dinge, Eigenschaften und Relationen, aus dem Russischen übersetzt von G. Kröber unter Mitarbeit von P. Bollhagen, Berlin 1965
- US Department of Housing and Urban Development (HUD), Urban and Regional Information Systems (URIS) - Support for Planning in Metropolitan; Teil 1, Washington, D.C. 1968

US Department of
Health, Education
and Welfare (HEW)

Towards a Social Report,
Washington, D.C. 1969

Überla, K.,

Faktorenanalyse - Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler, Berlin-Heidelberg-New York 1968

Vickery, B.C.,

Classification and indexing
in science, London 1959

Vischer, P.,

Simultane Produktions- und Absatzplanung, Bd. 3 der Schriftenreihe "Die Betriebswirtschaft in Forschung und Praxis", Hrsg. E. Heinen, Wiesbaden 1967

Wagner, H.,

Zum Problem der Zielfunktion in einer operationalen Theorie der betrieblichen Kapitaldisposition, Habilitationsschrift, Münster 1970

Wagner, H.,

Zielbildung und Entscheidungsprozeß bei multipolarer Struktur der betrieblichen Willensbildung, in: Sonderdruck aus: Schriften zur Kooperationsforschung, Theorie und Praxis der Kooperation, A. Studien, Bd. 3, Tübingen o.J., S. 91ff

Wahl, M.P.,

Grundlagen eines Management-Informationssystems: Einige wesentliche Definitionen, Sonderdruck aus IBM Nachrichten, 19. Jg., 1969

Webber, M.M.,

The Role of Intelligence Systems in Urban Planning, in: Journal of American Institute of Planners, November 1965, S. 289ff

- Weber, W./
Streißler, E., Erwartungen, Unsicherheit und Risiko, in: Handwörterbuch der Sozialwissenschaften, Bd. III, S. 330ff
- Wedekind, H., Datenorganisation, Berlin 1970
- Weichselberger, K., Genauigkeitsansprüche, Fehler und Kontrollen bei Volkszählungen, in: Allgemeines Statistisches Archiv, Bd. 54, 1970, S. 1ff
- Weizenbaum, J., ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine, in: Communications of the Association for Computing Machinery, New York 1966, S. 36ff
- Whorf, B.L., Sprache - Denken - Wirklichkeit: Beiträge zur Metalinguistik und Sprachphilosophie, hrsg. v. P. Krausser, Reinbek 1963
- Widdra, W., Einfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden I - Grundlagen und Versuchsplanung, hrsg. v. E. Walter, Berlin-Heidelberg-New York 1970, S. 324ff
- Widdra, W., Mehrfache Kovarianzanalyse, in: Lecture Notes in Operations Research and Mathematical Systems, Bd. 38: Statistische Methoden I - Grundlagen und Versuchsplanung, hrsg. v. E. Walter, Berlin-Heidelberg-New York 1970, S. 329ff
- Wild, J., Informationskostenrechnung auf der Grundlage informationeller Input-, Output- und Prozeßanalysen, in: Zeitschrift für betriebswirtschaftliche Forschung, 22. Jg., 1970, S. 218ff

- Wilde, D.J., Optimum Seek Methods, Englewood Cliffs, N.J. 1964
- Will, H.J., Betriebswirtschaftliche Informationssysteme - Versuch einer intelligenztechnischen Definition, in: Zeitschrift für betriebswirtschaftliche Forschung, 20. Jg., 1968, S. 648ff
- Will, H.J., Datenbank-Systeme, in: Zeitschrift für Betriebswirtschaft, 41. Jg., 1971, S. 815ff
- Wintgen, G., Zur mengentheoretischen Definition und Klassifizierung kybernetischer Systeme, in: Statistische Hefte, 11. Jg., 1970, S. 261ff
- Witte, E., Phasen-Theorem und Organisation komplexer Entscheidungsverläufe, in: Zeitschrift für betriebswirtschaftliche Forschung, 20. Jg., 1968, S. 625ff
- Wittmann, W., Unternehmung und unvollkommene Information - Unternehmerische Voraussicht - Ungewißheit und Planung, Köln und Opladen 1959
- Young, S., Organization as a Total System, in: Cleland, D.I./King, W.R., Hrg., Systems-Organizations-Analysis-Management, New York usw. 1969, S. 51ff
- Zangemeister, Ch., Marketing Informationssysteme: EDV-Einsatz im gesamten Absatzbereich, Bd. 5 der SCS Schriftenreihe, Hamburg o.J.

Lebenslauf

Ansgar Franz Breitung	457 Quakenbrück, An der Mühle 34
geboren am:	2. Januar 1945 in Quakenbrück
Eltern:	Josef Breitung, Finanzangestellter und Martha Breitung, geb. Trenkler
Schulbesuch:	1951-1955 kath. Volksschule in Quakenbrück 1955-1959 Artland-Gymnasium in Quakenbrück 1959-1962 Realschule in Quakenbrück 1962-1965 Gymnasium Meppen 1965 Reifeprüfung am Gymnasium Meppen
Studium:	1965-1970 Studium der Betriebswirt- schaftslehre in Münster 1970 Examen als Diplom-Kaufmann
Berufsweg:	ab 1970 Mitarbeiter am Institut für Marketing der Universität Münster



